

BACK TO THE FUTURE

Science in the Criminal Courtroom: Preventing Faulty Convictions

Temple American Inn of Court

November Team

November 8th, 2017

TABLE OF CONTENTS

1. Forensic Science: Daubert's Failure - Paul C. Giannelli
2. The (Near) Irrelevance of Daubert to Criminal Justice and Some Suggestions for Reform – Peter J. Neufield
3. Discovering Forensic Fraud – Jennifer D. Oliva & Valena E. Beety
4. Police Reports of Mock Suspect Interrogations: A Test of Accuracy and Perception – Saul M. Kassin, Jeff Kukucka, Victoria Z. Lawson, and John DeCarlo
5. Manipulation of Suspects and Unrecorded Questioning: After Fifty Years of Miranda Jurisprudence, Still Two (or Maybe Three) Burning Issues – Christopher Slobogin
6. Traces of Crime: How New York's DNA Techniques Became Tainted – Lauren Kirchner
7. Syndrome: Medical Uncertainty Casts Doubt on Convictions – Molly Gena
8. Junk Science and the Execution of an Innocent Man – Paul C. Giannelli
9. The Shifted Paradigm: Forensic Science's Overdue Evolution from Magic to Law – M. Chris Fabricant and Tucker Carrington

August 30, 2017

FORENSIC SCIENCE: *DAUBERT'S* FAILURE

59 CASE W. RES. L. REV. __ (forthcoming)

Paul C. Giannelli
Distinguish University Professor Emeritus
Case Western Reserve University

- I. Introduction
 - A. *Daubert*/Rule 702
 - B. National Academy of Sciences Forensic Report (2009)
- II. Forensic Techniques
 - A. Bite Mark Comparisons
 - B. Microscopic Hair Analysis
 - C. Firearms & Toolmark Identifications
 - D. Fingerprint Examinations
 - E. Comparative Bullet Lead Analysis
 - F. Arson Evidence
- III. Forensic Science Research
 - A. National Commission on Forensic Science (2013-17)
 - B. White House PCAST Report (2016)
- IV. Conclusion

“The man who discovers a new scientific truth has previously had to smash to atoms almost everything he had learnt, and arrives at the new truth with hands bloodstained from the slaughter of a thousand platitudes.” — Jose Oreta y Gasset, *The Revolt of the Masses* ch. XIV (1930).

I. INTRODUCTION

In 2015, Judge Alex Kozinski of the Ninth Circuit Court of Appeals noted that “[m]any defendants have been convicted and spent countless years in prison based on evidence by arson experts who were later shown to be little better than witch doctors.”¹ In the same year, Dr. Jo Handelsman, a White House science advisor, observed: “Suggesting that bite marks [should] still be a seriously used technology is not based on science, on measurement, on something that has standards, but more of a gut-level reaction.”² According to Judge Catherine Easterly of the D.C. Court of Appeals, “[a]s matters currently stand, a certainty statement regarding toolmark pattern matching has the same probative value as the vision of a psychic.”³ A New York Times editorial echoed these sentiments: “And the courts have only made the problem worse by purporting to be scientifically literate, and allowing in all kinds of evidence that would not make it within shouting distance of a peer-reviewed journal. Of the 329 exonerations based on DNA testing since 1989, more than one-quarter involved convictions based on ‘pattern’ evidence — like hair samples, ballistics, tire tracks, and bite marks — testified to by so-called experts.”⁴

¹ Alex Kozinski, *Criminal Law 2.0*, 44 GEO. L.J. ANN. REV. CRIM. PROC. iii, v (2015). See also *Almeciga v. Ctr. for Investigative Reporting, Inc.*, 185 F. Supp. 3d 401, 415 (S.D.N.Y. 2016) (“There have been too many pseudo-scientific disciplines that have since been exposed as profoundly flawed, unreliable, or baseless for any Court to take this [gate-keeping] role lightly.”).

² See Radley Balko, *A High-Ranking Obama Official Just Called for the “Eradication” of Bite Mark Evidence*, WASH. POST, July 22, 2015 (quoting remarks presented at the International Symposium on Forensic Science Error Management — Detection, Measurement and Mitigation, Arlington, Virginia (July 20–24, 2015), organized by the National Institute of Standards and Technology (NIST)).

³ *Williams v. United States*, 130 A.3d 343, 355 (D.C. 2016) (Easterly, J., concurring).

⁴ Editorial, *Junk Science at the F.B.I.*, N.Y. TIMES, Apr. 27, 2015. See also Eric S. Lander, *Fix the Flaws in Forensic Science*, N.Y. TIMES, Apr. 21, 2015 (“No expert should be permitted to testify without showing three things: a public database of patterns from many representative samples; precise and objective criteria for declaring matches; and peer-reviewed published studies that validate the methods.”).

These criticisms are valid — which raises a puzzling and consequential question: Why didn't the Supreme Court's "junk science" decision, *Daubert v. Merrell Dow Pharmaceuticals, Inc.*,⁵ prevent or restrict the admissibility of testimony based on flawed forensic techniques? *Daubert* was decided in 1993, nearly twenty-five years ago.

A. *Daubert/Rule 702*

Daubert was considered a revolutionary decision.⁶ It "radically changed the standard for admissibility of scientific testimony"⁷ by sweeping away the *Frye* "general acceptance" test,⁸ which had been the majority rule in both federal and state cases.⁹ The *Frye* standard gave great deference to the views of forensic practitioners and not to empirical testing.¹⁰ *Daubert* promised to be different. The Supreme Court held that "[p]roposed testimony must be supported by appropriate validation — i.e., 'good grounds,' based on what is known. In short, the requirement that an expert's testimony pertain to 'scientific knowledge' establishes a standard of evidentiary reliability."¹¹ In making this reliability determination, the *Daubert* Court highlighted five factors: (1) testing, (2) peer review and publication, (3) error rate, (4) maintenance of standards, and (5) general acceptance. The first and *most important factor* is empirical testing. The other factors are supplementary. Peer review and publication are designed to expose defects in testing. Acceptance of a technique within the scientific community is achieved through the publication of valid test results. Similarly, both error rates and standards are derived from testing.

Daubert was followed in 1999 by *Kumho Tire v. Carmichael*,¹² which held that *Daubert's* reliability standard applied to all expert testimony, not only

⁵ 509 U.S. 579 (1993).

⁶ See David L. Faigman, *Is Science Different for Lawyers?*, 297 SCIENCE 339, 340 (2002) ("*Daubert* initiated a scientific revolution in the law.>").

⁷ United States v. Barnette, 211 F.3d 803, 815 (4th Cir. 2000). See also United States v. Alatorre, 222 F.3d 1098, 1100 (9th Cir. 2000) ("*Daubert* has become ubiquitous in federal trial courts.>").

⁸ Frye v. United States, 293 F. 1013, 1014 (D.C. Cir. 1923) (stating that a technique "must be sufficiently established to have gained general acceptance in the particular field in which it belongs").

⁹ See Paul C. Giannelli, *The Admissibility of Novel Scientific Evidence: Frye v. United States, A Half-Century Later*, 80 COLUM. L. REV. 1197 (1980).

¹⁰ See Michael J. Saks, *Merlin and Solomon: Lessons from the Law's Formative Encounters with Forensic Identification Science*, 49 HASTINGS L.J. 1069, 1138 (1998) ("*Frye* does not work because its measure of validity is the judgment of 'the field,' and the field may consist of nonsense. For example, the *Frye* doctrine cannot exclude astrology.>").

¹¹ *Daubert*, 509 U.S. at 590 (emphasis added).

¹² 526 U.S. 137 (1999).

scientific evidence. By 2000, the Supreme Court was describing *Daubert* as establishing an “exacting” standard.¹³ In the same year, Federal Rule of Evidence 702 was amended to incorporate the *Daubert/Kumho* standard.¹⁴ Although a handful of jurisdictions continue to apply the *Frye* test, about forty jurisdictions have adopted the *Daubert* standard in one form or another.¹⁵

During this time, there was no shortage of commentary on the lack of empirical research in forensic science.¹⁶ For example, shortly after *Daubert* was decided, Professor Margaret Berger wrote: “Considerable forensic evidence made its way into the courtroom without empirical validation of the underlying theory and/or its particular application.”¹⁷ After *Kumho*, two commentators — citing bite mark, hair, and firearm analysis — observed that “little rigorous, systematic research has been done to validate the discipline’s basic premises and techniques, and in each area there was no evident reason why such research would be infeasible.”¹⁸

Notwithstanding *Daubert*’s promise, scholars soon discerned its uneven

¹³ Weisgram v. Marley Co., 528 U.S. 440, 455 (2000).

¹⁴ After *Daubert*, the Court decided *General Elec. Co. v. Joiner*, 522 U.S. 136 (1997), which established the standard for appellate review (abuse of discretion) for applying the *Daubert* factors. *Daubert*, *Joiner*, and *Kumho* make up what is known as the *Daubert* Trilogy.

¹⁵ See 1 PAUL C. GIANNELLI ET AL., SCIENTIFIC EVIDENCE § 1.06 (5th ed. 2012).

¹⁶ A few perceptive scholars had noted the lack of empirical testing prior to *Daubert*. See Randolph N. Jonakait, *Forensic Science: The Need for Regulation*, 4 HARV. J. L. & TECH. 109, 137 (1991) (“Forensic science is supported by almost no research. The laboratory practices are based on intuitions and deductions, not on empirical proof.”); D. Michael Risinger et al., *Exorcism of Ignorance as a Proxy For Rational Knowledge: The Lessons of Handwriting Identification “Expertise,”* 137 U. PA. L. REV. 731, 738 (1989) (“Our literature search for empirical evaluation of handwriting identification turned up one primitive and flawed validity study from nearly 50 years ago, one 1973 paper that raises the issue of consistency among examiners but presents only uncontrolled impressionistic and anecdotal information not qualifying as data in any rigorous sense, and a summary of one study in a 1978 government report. Beyond this, nothing.”); Michael J. Saks & Jonathan J. Koehler, *What DNA “Fingerprinting” Can Teach the Law About the Rest of Forensic Science*, 13 CARDOZO L. REV. 361, 372 (1991) (“[F]orensic scientists, like scientists in all other fields, should subject their claims to methodologically rigorous empirical tests. The results of these tests should be published and debated. Until such steps are taken, the strong claims of forensic scientists must be regarded with far more caution than they traditionally have been.”).

¹⁷ Margaret A. Berger, *Procedural Paradigms for Applying the Daubert Test*, 78 MINN. L. REV. 1345, 1354 (1994) (“Courts never required some of the most venerable branches of forensic science — such as fingerprinting, ballistics, and handwriting — to demonstrate their ability to make unique identifications.”).

¹⁸ Paul C. Giannelli & Edward J. Imwinkelried, *Scientific Evidence: The Fallout from the U.S. Supreme Court’s Decision in Kumho Tires*, 14 CRIMINAL JUSTICE, Winter 2000, at 12, 40. For an insightful analysis of how identification science was accepted by the courts, see Saks, *supra* note 10.

application in civil and criminal cases: “[T]he heightened standards of dependability imposed on expertise proffered in civil cases has continued to expand, but . . . expertise proffered by the prosecution in criminal cases has been largely insulated from any change in pre-*Daubert* standards or approach.”¹⁹ The title of a 2005 article pretty much summed up the state of the law — “*The (Near) Irrelevance of Daubert to Criminal Justice*.”²⁰ In short, *Daubert-lite*.

B. National Academy of Sciences Forensic Report (2009)

In 2006 Congress entered the picture by authorizing the National Academy of Sciences (NAS) to conduct a study of forensic science. After a three-year investigation, NAS issued a landmark report. One of its most riveting passages concluded: “Among existing forensic methods, only nuclear DNA analysis has been rigorously shown to have the capacity to consistently, and with a high degree of certainty, demonstrate a connection between an evidentiary sample and a specific individual or source.”²¹ The report went on to state that “some forensic science disciplines are supported by little rigorous systematic research to validate the discipline’s basic premises and techniques.”²² Such common forensic techniques as fingerprint examinations,²³ firearms (ballistics) and toolmark identifications,²⁴ handwriting examinations,²⁵ microscopic hair analysis,²⁶ and bite

¹⁹ D. Michael Risinger, *Navigating Expert Reliability: Are Criminal Standards of Certainty Being Left on the Dock?*, 64 ALB. L. REV. 99, 149 (2000). In addition, an extensive study of reported criminal cases found that “the *Daubert* decision did not impact on the admission rates of expert testimony at either the trial or appellate court levels.” Jennifer Groscup et al., *The Effects of Daubert on the Admissibility of Expert Testimony in State and Federal Criminal Cases*, 8 PYSCHOL. PUB. POL’Y & L. 339, 364 (2002).

²⁰ Peter J. Neufeld, *The (Near) Irrelevance of Daubert to Criminal Justice and Some Suggestions for Reform*, 95 AM. J. PUB. HEALTH S107 (2005).

²¹ NATIONAL RESEARCH COUNCIL, NATIONAL ACADEMY OF SCIENCES, STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD 100 (2009) [hereinafter NAS FORENSIC REPORT].

²² *Id.* at 22. At another point, the report stated: “The simple reality is that the interpretation of forensic evidence is not always based on scientific studies to determine its validity. This is a serious problem.” *Id.* at 8. See also *id.* at 6 (“Often there are no standard protocols governing forensic practice in a given discipline. And, even when protocols are in place . . . , they often are vague and not enforced in any meaningful way.”).

²³ *Id.* at 144 (Research is needed “[t]o properly underpin the process of friction ridge [fingerprint] identification.”).

²⁴ *Id.* at 154 (“Sufficient studies [on firearms identification] have not been done to understand the reliability and repeatability of the methods.”).

²⁵ *Id.* at 166 (“The scientific basis for handwriting comparisons needs to be strengthened.”).

²⁶ *Id.* at 161 (“[T]estimony linking microscopic hair analysis with particular defendants is highly unreliable.”).

mark comparisons²⁷ fell into this category.

Not only did the NAS report highlight flaws in forensic science, it sharply criticized the judiciary for failing to demand the validation that *Daubert* required: “The bottom line is simple: In a number of forensic science disciplines, forensic science professionals have yet to establish either the validity of their approach or the accuracy of their conclusions, *and the courts have been utterly ineffective in addressing this problem.*”²⁸ In a later passage, the report declared that “*Daubert* has done little to improve the use of forensic science evidence in criminal cases.”²⁹ The disparate treatment of civil actions and criminal prosecutions was also noted. After finding that “trial judges rarely exclude or restrict expert testimony offered by prosecutors,” the report commented: “[I]ronically, the appellate courts appear to be more willing to second-guess trial court judgments on the admissibility of purported scientific evidence in civil cases than in criminal cases.”³⁰

Despite the NAS report, courts generally continued to admit the same evidence. Only a handful of courts applied the “exacting” standard that the Supreme Court said *Daubert* demanded.³¹

* * *

This article examines the justice system’s failure by reviewing the status of six forensic techniques: (1) bite mark analysis, (2) microscopic hair comparisons, (3) firearms and toolmark identifications, (4) fingerprint examinations, (5) bullet lead analysis, and (6) arson evidence. It argues that the system’s failure can be traced back to its inability to demand and properly evaluate foundational research, i.e., *Daubert*’s first factor (empirical testing). Indeed, the justice system may be structurally incapable of applying *Daubert* in criminal cases.

A different paradigm is needed, one that assigns an independent agency the responsibility of evaluating foundational research. As discussed in Part III, this approach was recently recommended by the National Commission on

²⁷ *Id.* at 174 (“No thorough study has been conducted of large populations to establish the uniqueness of bite marks . . .”).

²⁸ *Id.* at 53 (emphasis added).

²⁹ *Id.* at 106.

³⁰ *Id.* at 11.

³¹ *Weisgram*, 528 U.S. at 455. As former federal district judge Nancy Gertner noted: “[A] busy trial judge can rely on the decades of case law to legitimize decisions rejecting a hearing or motions in limine. And the trial judge can count on the Court of Appeals likely concluding that rejecting the challenge was not an abuse of the judge’s discretion.” Nancy Gertner, *Commentary on the Need for A Research Culture in the Forensic Sciences*, 58 UCLA L. REV. 789, 790 (2011).

Forensic Science (2013-17)³² and the President’s Council of Advisors on Science and Technology (2016) (PCAST).³³ Both recommended that the National Institute of Standards and Technology (NIST) evaluate all forensic disciplines on a continuing basis, thereby injecting much needed scientific expertise into the criminal justice system.

II. FORENSIC TECHNIQUES

A. Bite Mark Comparisons

For decades, bite mark evidence has been admitted in hundreds of trials,³⁴ many of which were capital prosecutions.³⁵ No reported American case has rejected bite mark testimony. Moreover, it is not uncommon for courts to speak of bite mark comparison as a “science”³⁶ — even an “exact science.”³⁷ Acceptance of the technique is so deeply entrenched that some courts have taken judicial notice of its validity,³⁸ which means its reliability is indisputable.³⁹ Distinctive characteristics of a person’s dentition were first used to identify

³² In 2013, the Department of Justice (DOJ), in partnership with the National Institute of Standards and Technology (NIST), established the National Commission on Forensic Science to enhance the practice and improve the reliability of forensic science. The author served on the Commission.

³³ PRESIDENT’S COUNCIL OF ADVISORS ON SCIENCE AND TECHNOLOGY, REPORT TO THE PRESIDENT: FORENSIC SCIENCE IN CRIMINAL COURTS: ENSURING SCIENTIFIC VALIDITY OF FEATURE-COMPARISON METHODS (2016) [hereinafter WHITE HOUSE PCAST Report].

³⁴ See 1 GIANNELLI ET AL., *supra* note 15, § 13.05 (discussing the admissibility of bite mark evidence). In *Doyle v. State*, 263 S.W.2d 779 (Tex. Crim. App. 1954), a bite mark was left in a piece of cheese in a burglary case. Two decades later, in *Patterson v. State*, 509 S.W.2d 857, 862 (Tex. Crim. App. 1974), a prosecution expert matched the defendant’s teeth to a mark found on a murder victim.

³⁵ See Carrington Tucker, *Mississippi Innocence: The Convictions and Exonerations of Leven Brooks and Kennedy Brewer and the Failure of the American Promise*, 28 GEO. J. LEGAL ETHICS 123 (2015).

³⁶ See *People v. Marsh*, 441 N.W.2d 33, 35 (Mich. Ct. App. 1989) (“the science of bite mark analysis has been extensively reviewed in other jurisdictions”).

³⁷ See *State v. Sager*, 600 S.W.2d 541, 569 (Mo. Ct. App. 1980) (“an exact science”).

³⁸ See *State v. Richards*, 804 P.2d 109, 112 (Ariz. 1990) (“[B]ite mark evidence is admissible without a preliminary determination of reliability . . .”); *People v. Middleton*, 429 N.E.2d 100, 101 (N.Y. 1981) (“The reliability of bite mark evidence as a means of identification is sufficiently established in the scientific community to make such evidence admissible in a criminal case, without separately establishing scientific reliability in each case . . .”); *State v. Armstrong*, 369 S.E.2d 870, 877 (W. Va. 1988) (judicially noticing the reliability of bite mark evidence).

³⁹ See FED. R. EVID. 201(b) (limiting judicial notice to a “fact that is not subject to reasonable dispute”).

skeletonized remains and individuals in mass disasters such as plane crashes.⁴⁰ Courts assumed that these distinctive characteristics can be transferred to another person's skin during a violent crime (e.g., homicides, rapes, and child abuse),⁴¹ an assumption that overlooked some obvious problems. First, bite marks typically involve no more than the edges of six to eight front teeth, not thirty-two teeth with five anatomical surfaces that can be used when comparing a deceased person's dentition with X-rays. Second, bite marks do not reveal artifacts such as fillings, crowns, etc., all of which assist in associating human remains with a person's dental records.⁴² Moreover, human skin is extremely malleable and thus subject to various types of distortion.⁴³ In addition, bite mark analysis is a subjective technique with no agreed-upon methodology.

1. *Foundational Research*

Despite overwhelming judicial approval, bite mark evidence is not supported by foundational research.⁴⁴ Indeed, the only rigorous studies are recent

⁴⁰ 1 GIANNELLI ET AL., *supra* note 15, at § 13.03 (discussing the admissibility of dental identifications).

⁴¹ *See* *People v. Milone*, 356 N.E.2d 1350, 1358 (Ill. App. Ct. 1976) (“The concept of identifying a suspect by matching his dentition to a bite mark found at the scene of a crime is a logical extension of the accepted principle that each person’s dentition is unique.”); *People v. Smith*, 443 N.Y.S.2d 551, 556-57 (Cty. Ct. 1981) (“The basic premise is the unique nature of individual dentition ... and the virtually infinite number of individual bite configurations.”).

⁴² “Restorations alone, with varying shapes, sizes, and restorative materials, may offer numerous points for comparison. In addition to restorations, the number of teeth, prostheses, decay, malposition, malrotation, peculiar shapes, root canal therapy, bone patterns, bite relationship, and oral pathology may all provide identifying characteristics.” 1 GIANNELLI ET AL., *supra* note 15, at 711.

⁴³ *See* I.A. Pretty & D. Sweet, *The Scientific Basis for Human Bite Mark Analyses—A Critical Review*, 41 SCI. & JUST. 85, 87 (2001) (“Skin is a poor registration material since it is highly variable in terms of anatomical location, underlying musculature or fat, curvature, and looseness or adherence to underlying tissues. Skin is highly visco-elastic, which allows stretching to occur during either the biting process or when evidence is collected.”).

One study classified different types of distortion: Primary distortion occurs at the time of biting and results (1) from the dynamics of the biting process (dynamic distortion) and (2) from the features of the tissue bitten (tissue distortion). Secondary distortion occurs at a subsequent time. It can be subdivided into three categories. The first is time-related distortion, e.g., caused by subsequent healing or decomposition. Posture distortion results when the bite mark is viewed or recorded in a position that differs from the position at the time of biting. Photographic distortion results from the angle of the camera and the curvature of the body. *See* D.R. Sheasby & D.G. MacDonald, *A Forensic Classification of Distortion in Human Bite Marks*, 122 FORENSIC SCI. INT’L 75 (2001).

⁴⁴ *See* Saks, *supra* note 10, at 1120 (“[R]ather than the field convincing the courts of the sufficiency of its knowledge and skills, admission by the courts apparently convinced the forensic odontology community that, despite their doubts, they really were able to perform bite mark identifications.”).

— and undercut the technique’s validity.⁴⁵ The 2009 NAS forensic report concluded that “the scientific basis is insufficient to conclude that bite mark comparisons can result in a conclusive match.”⁴⁶ Despite the NAS report, courts continued to permit expert testimony on the subject. For example, in *State v. Prade*,⁴⁷ decided in 2014, the expert testified that “bite mark evidence is generally accepted within the scientific community.”⁴⁸ Similarly, in *Coronado v. State*,⁴⁹ a different expert stated that he did not “agree with the NAS Report’s conclusion that bite mark analysis cannot result in a conclusive match” — adding “you do not have to be a ‘rocket scientist’ to see that, in some cases, there is a unique and distinct pattern of teeth that can be identified.”⁵⁰ In addition, these experts rejected the valid research mentioned above⁵¹ and both prosecutors and their

⁴⁵ Dr. Mary Bush and her colleagues at the Laboratory for Forensic Odontology, State University of New York at Buffalo, have published over a dozen studies that have undermined the assumptions underpinning bite mark evidence. See, e.g., Mary Bush et al., *Statistical Evidence for the Similarity of the Human Dentition*, 56 J. FORENSIC SCI. 118, 122 (2011) (“Our results show that given our measurement parameters, statements concerning dental uniqueness with respect to bitemark analysis in an open population are unsupported. . . . Confidence in the notion of dental uniqueness in bitemark analysis has been based on anecdotal knowledge, the use of inappropriate statistics, and precedence of admission in the courtroom.”); Mary Bush et al., *Biomechanical Factors in Human Dermal Bitemarks in a Cadaver Model*, 54 J. FORENSIC SCI. 167 (2009) (23 bites were made in cadaver skin with the same dentition using an instrumented-biting machine. The cadavers were moved and re-photographed in different positions. Subsequent measurements showed differences between all bite marks. In addition, postural distortion was significant).

One survey of fifteen odontologists involved their opinions of six images of supposed bite marks. The “practitioner agreement was at best fair, with wide-ranging opinions on the origin, circumstance, and characteristics of the wound given for all six images.” M. Page et al., *Expert Interpretation of Bitemark Injuries — A Contemporary Qualitative Study*, 58 J. FORENSIC SCI. 664, 664 (2013).

⁴⁶ NAS FORENSIC REPORT, *supra* note 21, at 175.

⁴⁷ 9 N.E.3d 1072 (Ohio Ct. App. 2014).

⁴⁸ *Id.* at 1097.

⁴⁹ 384 S.W.3d 919 (Tex. Ct. App. 2012).

⁵⁰ *Id.* at 926.

⁵¹ See *Prade*, 9 N.E.3d at 1098 (“As to Dr. Bush’s cadaver studies, Dr. Wright testified that cadaver skin simply cannot compare with living skin. Dr. Wright explained that cadaver skin only distorts after a bite for two to three minutes at most because, unlike live skin, no bruising, contusions, or lacerations occur. Dr. Wright also testified that using a mechanical jaw to bite is problematic because the jaw operates on a fixed hinge that cannot mimic the wider range of movement that an actual jaw is capable of.”). But see I.A. Pretty & D. Sweet, *A Paradigm Shift in the Analysis of Bitemarks*, 201 FORENSIC SCI. INT’L 38, 40 (2010) (cadaver models have limitations but “there is little alternative for researchers to produce bitemarks of known origin”; use of anesthetized pigs to create peri-mortem injuries raises a different issue — i.e., differences between pigskin and human skin).

experts attacked researchers without offering any foundational research.⁵²

Unfortunately, the American Board of Forensic Odontology (ABFO) has fiercely defended bite mark analysis. To bolster its position, the ABFO conducted a study that was presented at a forensic conference in 2015.⁵³ As it turned out, the study undercut the ABFO's own position. Thirty-nine ABFO-certified bite mark experts — with an average of twenty years experience — examined 100 bite mark photographs. Each was asked three questions:

(1) Is there sufficient evidence in the presented materials to render an opinion on whether the patterned injury is a human bite mark?

(2) Is it a human bite mark, not a human bite mark, or suggestive of a human bite mark?

(3) Does the bite mark have distinct, identifiable arches and individual tooth marks?

The results to the first question were not reassuring. The thirty-nine experts agreed unanimously in only four out of the 100 cases. In only twenty cases was there 90 or more percent agreement. At the end of question two — whether the mark is a human bite mark — there were only sixteen cases with 90 or more percent agreement. At the end of the third question, there were only eight cases in which at least 90 percent of the analysts agreed.⁵⁴ Equally disturbing was the ABFO's decision to postpone publishing the results “until the organization can tweak the design of the study and conduct it again, a process that's expected to take at least a year.”⁵⁵ In effect, a do-over. Meanwhile, an Associated Press analysis reported that at least twenty-four men convicted or charged with murder or rape based on bite marks have been exonerated since 2000.⁵⁶

2. *Texas Forensic Science Commission (2016)*

Steven Chaney spent twenty-eight years in prison for murder based largely

⁵² See Radley Balko, *In Angry, Defensive Memo, Manhattan DA's Office Withdraws Bite Mark Evidence*, WASH. POST., Jan. 13, 2016; Radley Balko, *Attack of the Bite Mark Matchers*, WASH. POST, Feb. 18, 2015; Radley Balko, *The Path Forward on Bite Mark Matching — and the Rearview Mirror*, WASH. POST, Feb. 20, 2015.

⁵³ The study is known as Construct Validity Bitemark Assessments Using the ABFO Bitemark Decision Tree (“Freeman/Pretty Study”).

⁵⁴ Radley Balko, *A Bite Mark Matching Advocacy Group Just Conducted a Study that Discredits Bite Mark Evidence*, WASH. POST, Apr. 8, 2015.

⁵⁵ *Id.*

⁵⁶ See *Chaney v. State*, 775 S.W.2d 722, 725 (Tex. App. 1989) (A board-certified forensic odontologist “concluded that, in his opinion and with reasonable dental certainty, appellant made the bite marks on Sweek's body.”).

on bite mark evidence. When he was eventually exonerated via DNA testing,⁵⁷ the Innocence Project filed a complaint on his behalf with the Texas Forensic Science Commission (TFSC).⁵⁸ In 2016, after a six-month investigation, the TFSC recommended a moratorium on the admission of bite mark testimony. It found that there is no scientific basis for claiming that a particular mark can be associated to a person's dentition: "Any testimony describing human dentition as 'like a fingerprint' or incorporating similar analogies lacks scientific support."⁵⁹ Similarly, "there is no scientific basis for assigning probability or statistical weight to an association, regardless of whether such probability or weight is expressed numerically (e.g., 1 in a million) or using some form of verbal scale (e.g., highly likely/unlikely)."⁶⁰

TFSC was also alarmed that the ABFO study was not published due to "political and organizational pressures." In the commission's view, "such a resistance to publish scientific data contradicts the ethical and professional obligations of the profession as a whole, and is especially disconcerting when one considers the life and liberty interests at stake in criminal cases."⁶¹

3. *White House PCAST Report (2016)*

In September 2016, the White House released its report on forensic science.⁶² Regarding bite mark analysis, it concluded that (1) appropriately designed validation studies are lacking, (2) the few available studies had "very high" false-positive rates, (3) "inappropriate closed-set designs . . . are likely to underestimate the true false positive rate," and (4) the studies show that experts "cannot even consistently agree on whether an injury *is* a human bitemark."⁶³

⁵⁷ Amanda Lee Myers, *Once Key in Some Cases, Bite-mark Evidence Now Derided as Unreliable*, Associated Press, Jun. 16, 2013.

⁵⁸ Texas created the Texas Forensic Science Commission (TFSC) in 2005 after a scandal required Houston to close its crime lab. See TEX. CRIM. PROC. CODE art. 38.01(4)(a)(3) (2005) (among other duties, the Commission should "investigate, in a timely manner, any allegation of professional negligence or misconduct that would substantially affect the integrity of the results of a forensic analysis conducted by an accredited laboratory, facility, or entity"). See generally Michael Hall, *False Impressions*, TEXAS MONTHLY, Jan. 2016.

⁵⁹ TEXAS FORENSIC SCI. COMM'N, FORENSIC BITEMARK COMPARISON COMPLAINT FILED BY NATIONAL INNOCENCE PROJECT ON BEHALF OF STEVEN MARK CHANEY—FINAL REPORT 11-12 (2016).

⁶⁰ *Id.* at 12.

⁶¹ *Id.* at 13. See also Brandi Grissom, *Arguments Over Bitemarks Get Testy at Texas Forensic Science Commission Meeting*, DALLAS MORNING NEWS, Nov. 17, 2015.

⁶² WHITE HOUSE PCAST REPORT, *supra* note 33, at 9.

⁶³ "PCAST finds that bitemark analysis is far from meeting the scientific standards for foundational validity." *Id.*

Numerous cases support the last observation.⁶⁴

* * *

In sum, the courts have yet to reject bite mark evidence — a subjective method that is not supported by foundational research and lacks agreed-upon standards.⁶⁵ “Perhaps no discredited forensic assay has benefitted more from criminal courts’ abdication of gatekeeper responsibilities than bite mark analysis.”⁶⁶ Instead, it was the Innocence Project that spearheaded the challenges in this area, and in 2016 the Texas Forensic Science Commission became the first governmental body to seriously scrutinize the technique. Notwithstanding the NAS, PCAST, and TFSC reports, courts continue to admit bite mark evidence.⁶⁷ Incredibly, at the April 11, 2017 meeting of the National Commission on Forensic Science the chairman of the National District Attorneys Association stated that his organization believes that bite mark evidence is a “reliable science.”⁶⁸ A day earlier, Keith Harward described how bite mark evidence resulted in thirty-three

⁶⁴ See, e.g., *Ege v. Yukins*, 380 F. Supp. 2d 852, 878 (E.D. Mich. 2005) (“[T]he defense attempted to rebut Dr. Warnick’s testimony with the testimony of other experts who opined that the mark on the victim’s cheek was the result of *livor mortis* and was not a bite mark at all.”); *Czapleski v. Woodward*, 1991 U.S. Dist. Lexis 12567 (N.D. Cal. Aug. 30, 1991) (dentist’s initial report concluded that “bite” marks found on child were consistent with dental impressions of mother; several experts later established that the marks on child’s body were postmortem abrasion marks and not bite marks); *Kinney v. State*, 868 S.W.2d 463, 468 (Ark. 1994) (disagreement that marks were human bite marks); *People v. Noguera*, 842 P.2d 1160, 1165 n.1 (Cal. 1992) (“At trial, extensive testimony by forensic odontologists was presented by both sides, pro and con, as to whether the wounds were human bite marks and, if so, when they were inflicted.”); *State v. Duncan*, 802 So. 2d 533, 553 (La. 2001) (“Both defense experts testified that these marks on the victim’s body were not bite marks.”); *Stubbs v. State*, 845 So. 2d 656, 668 (Miss. 2003) (“Dr. Galvez denied the impressions found on Williams were the results of bite marks.”).

⁶⁵ See Michael J. Saks et al., *The Impending Death of Forensic Bitemark Identification*, 3 J. L. & BIOSCIENCES 1 (2016) (“[R]ecent reviews of the field’s claims, as well as recent empirical findings, have underscored the lack of reliability and validity of the most fundamental claims about the ability of forensic dentists to identify the source of bitemarks on human skin.”).

⁶⁶ M. Chris Fabricant & William Tucker Carrington, *The Shifted Paradigm: Forensic Science’s Overdue Evolution from Magic to Law*, 4 VA. J. CRIM. L. 1, 38 (2016).

⁶⁷ In *Commonwealth v. Ross*, No. CR 2038-2004, at 5 (C.P. Blair County, Pa., Mar. 8, 2017), the court admitted bite mark evidence, albeit limited, noting that “[t]he Commonwealth notes that no state or federal court has suppressed expert testimony in a criminal case based upon the NAS Report, and no courts have prohibited bite mark evidence based upon the PCAST or TFSC reports.” See also Radley Balko, *Incredibly, Prosecutors are Still Defending Bite Mark Evidence*, WASH. POST, Jan 30, 2017.

⁶⁸ Pem Levy, *Sessions’ New Forensic Science Adviser Has a History of Opposing Pro-Science Reforms*, MOTHER JONES, Aug. 10, 2017.

years of imprisonment before being exonerated by DNA evidence.⁶⁹

B. Microscopic Hair Analysis

In this examination, samples are first examined to identify features visible to the naked eye such as color and form, i.e., whether it is straight, wavy, or curved. Next, the sample is viewed microscopically to determine characteristics such as shaft form, hair diameter, and pigment size.

Experts have long acknowledged that a positive identification is not possible with microscopic hair analysis. Instead, examiners testify that a crime scene exemplar was “consistent with” a hair sample from the defendant. The probative value of this conclusion would, of course, vary if only a hundred people had microscopically indistinguishable hair as opposed to several million. Due to a lack of research, no one knows whether the crime scene hair could have come from 10 other persons or 100, 10,000, and so forth.⁷⁰ This important qualifying information was often omitted from the experts’ testimony, thus making marginal evidence appear misleadingly convincing.⁷¹

However, experts frequently went way beyond the “consistent with” language in their testimony, suggesting a rare association. For example, in the Edward Honaker case, the expert testified that the crime scene hair sample “was unlikely to match anyone” other than the defendant.⁷² Honaker spent ten years in

⁶⁹ See Frank Green, *DNA Proves Man Innocent of 1982 Rape and Murder in Famous ‘Bite-mark’ Case, Lawyers Say*, RICHMOND TIMES-DISPATCH, Mar. 12, 2016 (Keith Harward case); Spencer S. Hsu, *Va. Exoneration Underscores Mounting Challenges to Bite-Mark Evidence*, WASH. POST, Apr. 8, 2016 (Keith Harward case).

⁷⁰ As one hair examiner wrote: “If a pubic hair from the scene of a crime is found to be similar to those from a known source, [the courts] do not know whether the chances that it could have originated from another source are one in two or one in a billion.” B.D. Gaudette, *Probabilities and Human Pubic Hair Comparisons*, 21 J. FORENSIC SCI. 514, 514 (1976).

⁷¹ Professor Berger explained the problem:

We allow eyewitnesses to testify that the person fleeing the scene wore a yellow jacket and permit proof that a defendant owned a yellow jacket without establishing the background rate of yellow jackets in the community. Jurors understand, however, that others than the accused own yellow jackets. When experts testify about samples matching in every respect, the jurors may be oblivious to the probability concerns if no background rate is offered, or may be unduly prejudiced or confused if the probability of a match is confused with the probability of guilt, or if a background rate is offered that does not have an adequate scientific foundation.

Berger, *supra* note 17, at 1357.

⁷² EDWARD CONNORS ET AL., CONVICTED BY JURIES, EXONERATED BY SCIENCE: CASE STUDIES IN THE USE OF DNA EVIDENCE TO ESTABLISH INNOCENCE AFTER TRIAL 58 (1996).

prison before DNA proved him innocent.⁷³ In another case, an expert testified that hair samples were “consistent microscopically” but then elaborated: “In other words, hairs are not an absolute identification, but they either came from this individual or there is — could be another individual *somewhere in the world* that would have the same characteristics to their hair.”⁷⁴ This is an implicit (and extreme) probability statement that lacks any empirical support.

Although microscopic hair analysis had long been judicially accepted,⁷⁵ its validity was suspect.⁷⁶ In 1995, a federal district court in *Williamson v. Reynolds* observed: “Although the hair expert may have followed procedures accepted in the community of hair experts, the human hair comparison results in this case were, nonetheless, scientifically unreliable.”⁷⁷ The court also noted that the “expert did not explain which of the ‘approximately’ 25 characteristics were consistent, any standards for determining whether the samples were consistent, how many persons could be expected to share this same combination of characteristics, or how he arrived at his conclusions.”⁷⁸ *Williamson*, who was five days from execution when he obtained habeas relief, was subsequently exonerated by DNA testing.⁷⁹

The *Williamson* opinion — perhaps the only thorough judicial analysis of microscopic hair comparisons — was all but ignored by other courts. In *Johnson v. Commonwealth*⁸⁰ (1999), the Kentucky Supreme Court upheld the admissibility of hair evidence, taking “judicial notice” of its reliability⁸¹ and thus implicitly

⁷³ *Id.*

⁷⁴ *Williamson v. Reynolds*, 904 F. Supp. 1529, 1554 (E.D. Okl. 1995) (emphasis added), *rev'd on this issue sub nom.*, *Williamson v. Ward*, 110 F.3d 1508, 1523 (10th Cir. 1997) (holding that due process, not *Daubert*, controls in federal habeas review).

⁷⁵ See Edward J. Imwinkelried, *Forensic Hair Analysis: The Case Against the Underemployment of Scientific Evidence*, 39 WASH. & LEE L. REV. 41, 62 (1982) (stating that “[t]he massive body of case law, liberally admitting even hair evidence of low probative value, dwarfs the handful of cases excluding hair evidence”).

⁷⁶ See Clive A. Stafford Smith & Patrick D. Goodman, *Forensic Hair Comparison Analysis: Nineteenth Century Science or Twentieth Century Snake Oil?*, 27 COLUM. HUM. RTS. L. REV. 227, 231 (1996) (“If the purveyors of this dubious science cannot do a better job of validating hair analysis than they have done so far, forensic hair comparison analysis should be excluded altogether from criminal trials.”).

⁷⁷ *Williamson*, 904 F. Supp. at 1558.

⁷⁸ *Id.* at 1554.

⁷⁹ See BARRY SCHECK ET AL., ACTUAL INNOCENCE: FIVE DAYS TO EXECUTION AND OTHER DISPATCHES FROM THE WRONGLY CONVICTED 146 (2000) (noting that the hair evidence was shown to be “patently unreliable”). See also JOHN GRISHAM, THE INNOCENT MAN: MURDER AND INJUSTICE IN A SMALL TOWN (2006) (examining *Williamson*’s trial).

⁸⁰ 12 S.W.3d 258 (Ky. 1999).

⁸¹ *Id.* at 267.

finding its validity indisputable.⁸² Other courts echoed *Johnson*, not *Williamson*.⁸³ Indeed, ten years after *Williamson* was decided, a 2005 decision by the Connecticut Supreme Court observed (correctly) that “[t]he overwhelming majority of courts have deemed such evidence admissible.”⁸⁴

Once again, the courts abdicated their responsibility. Indeed, hair evidence only began to be carefully scrutinized after a startling number of DNA exonerations were reported.⁸⁵ A 2008 study of 200 DNA exonerations found that expert testimony (55 percent) was the second leading type of evidence — after eyewitness identifications (79 percent) — used in wrongful conviction cases.⁸⁶ A subsequent investigation of trial transcripts underscored the role of hair analysis in the exoneration cases: “Of the 65 cases involving microscopic hair comparison in which transcripts were located, 25 cases, or 38%, had invalid forensic science testimony.”⁸⁷ The 2009 NAS report observed that “testimony linking microscopic hair analysis with particular defendants is highly unreliable.”⁸⁸

1. *FBI Hair Review*

In May 2013, the Mississippi Supreme Court, in a 5-to-4 decision, rejected Willie Jerome Manning’s request for a stay of execution to permit DNA testing — “potentially setting up what experts said would be a rare case in recent years in which a person is put to death with such requests unmet.”⁸⁹ A week later, the court unexpectedly stayed Manning’s execution — after the Department of Justice (DOJ) notified state officials that FBI experts had presented misleading testimony

⁸² See FED. R. EVID. 201(b) (limiting judicial notice to a “fact that is not subject to reasonable dispute”).

⁸³ See 2 GIANNELLI ET AL., *supra* note 15, § 24.03, at 825 (noting the “limited impact of *Daubert*”).

⁸⁴ State v. West, 877 A.2d 787, 808 (Conn. 2005).

⁸⁵ In 1998, a Canadian judicial inquiry into the wrongful conviction of Guy Paul Morin was released. Morin’s original conviction was based, in part, on hair evidence. The judge conducting the inquiry recommended that “[t]rial judges should undertake a more critical analysis of the admissibility of hair comparison evidence as circumstantial evidence of guilt.” HON. FRED KAUFMAN, THE COMMISSION ON PROCEEDINGS INVOLVING GUY PAUL MORIN (Ontario Ministry of the Attorney General 1998) (Recommendation 2). See also EDWARD CONNORS ET AL., *supra* note 72, 58 (listing cases).

⁸⁶ Brandon L. Garrett, *Judging Innocence*, 108 COLUM. L. REV. 55, 81 (2008).

⁸⁷ Brandon L. Garrett & Peter Neufeld, *Invalid Forensic Science Testimony and Wrongful Convictions*, 95 VA. L. REV. 1, 14-15 (2009).

⁸⁸ NAS FORENSIC REPORT, *supra* note 21, at 161.

⁸⁹ Campell Robertson, *Mississippi Inmate’s Bid for DNA Tests Is Denied With Tuesday Execution Set*, N.Y. TIMES, May 4, 2013, at A11.

at his trial, including hair and firearms evidence.⁹⁰

Soon after, the DOJ announced that Manning was but one of 120 cases — including twenty-seven death penalty prosecutions — in which improper microscopic hair analysis had been introduced in evidence.⁹¹ For example, examiners claimed to connect a hair sample to a single person “to the exclusion of all others” or stated or suggested a probability for such a match from past casework.⁹² The FBI review came after three District of Columbia men, who had been convicted of rape or murder in the early 1980s, were exonerated through DNA testing.⁹³ In one of these cases, the FBI expert testified: “Chances that it came from someone else were ‘one in 10 million.’”⁹⁴

After further investigation, DOJ reported in 2015 that “FBI examiners had provided scientifically invalid testimony in more than 95 percent of cases where that testimony was used to inculcate a defendant at trial.”⁹⁵ *Commonwealth v. Perrott*⁹⁶ was one of the first cases to reach the courtroom as a consequence of the DOJ review. A superior court granted Perrott a new trial in 2016, criticizing the misleading use of hair evidence. The court noted: “In discussing the ‘microscopic characteristics’ of hair, [the expert] stated that these characteristics ‘make that hair somewhat unique.’ He likened the ‘subtle’ characteristics of hair that ‘make it somewhat unique’ to the subtle differences in a human face.”⁹⁷

⁹⁰ See Campell Robertson, *With Hours Left to Go, Execution Is Postponed*, N.Y. TIMES, May 8, 2013 (noting that the Department of Justice “disavow[ed] the degree of certainty expressed by F.B.I. forensic experts at the man’s trial”). See also Andrew Cohen, *A Ghost of Mississippi: The Willie Manning Capital Case*, THE ATLANTIC, May 2, 2013.

⁹¹ See Jack Nicas, *Flawed Evidence Under a Microscope: Disputed Forensic Techniques Draw Fresh Scrutiny; FBI Says It Is Reviewing Thousands of Convictions*, WALL ST. J., July 18, 2013.

⁹² Spencer S. Hsu, *U.S. Reviewing 27 Death Penalty Convictions for FBI Forensic Testimony Errors*, WASH. POST, July 17, 2013 (“[O]n the witness stand, several agents for years went beyond the science and testified that their hair analysis was a near-certain match.”).

⁹³ See Editorial, *Failures at the FBI Crime Lab*, WASH. POST, Apr. 20, 2012 (“Kirk L. Odom was incarcerated for 20 years and Donald E. Gates for nearly 30 for crimes they did not commit. Santae A. Tribble spent 28 years behind bars, even though DNA evidence now shows he almost undoubtedly was not the culprit.”).

⁹⁴ Martin Enserink, *Evidence on Trial*, 351 SCIENCE 1129, 1129, Mar. 11, 2016.

⁹⁵ WHITE HOUSE PCAST REPORT, *supra* note 33 at 3. See also Editorial, *Junk Science at the F.B.I.*, N.Y. TIMES, Apr. 27, 2015 (“a sweeping post-conviction review of 2,500 cases”); Hugh B. Kaplan, *DOJ Examiners Gave Bad Testimony in 90 Percent of Hair Comparison Cases*, BNA CRIM. L. RPTR. 77, Apr. 22, 2015.

⁹⁶ Nos. 85–5415, 5416, 5418, 5420, 5425, 2016 WL 380123 (Mass. Super. Ct. Jan. 26, 2016).

⁹⁷ *Id.* at *32. The expert also “asserted that the hairs ‘matched’ and showed a ‘strong association.’ In discussing the chance that the hair found on the victim’s bed came from someone other than Perrot, [the expert] conceded the possibility, adding that during his ten years

2. White House PCAST Report (2016)

In June 2016, the Department of Justice released proposed guidelines concerning hair testimony. Documentation purporting to support the validity and reliability of hair evidence accompanied the guidelines.⁹⁸ Listing several studies, the FBI concluded:

Based on these and other published studies, microscopic hair comparison has been demonstrated to be a valid and reliable scientific methodology. These studies have also shown that microscopic hair comparisons alone cannot lead to personal identification and it is crucial that this limitation be conveyed both in the written report and in testimony.⁹⁹

The White House PCAST report, however, challenged the supporting documentation, which discussed only a handful of studies from the 1970s and 1980s but did not comment on subsequent studies that found “substantial flaws in the methodology and results of the key papers.”¹⁰⁰ Moreover, “PCAST’s own review of the cited papers [found] that these studies do not establish the foundational validity and reliability of hair analysis.”¹⁰¹

* * *

The bottom line, again, is the judiciary’s dereliction in failing to curb the misuse of hair microscopy testimony. The Innocence Project’s track record of DNA exonerations brought this issue to the fore. Indeed, the three exonerations in the District of Columbia triggered the FBI review. Yet, DOJ’s proposed

of experience ‘it’s extremely rare that I will have known hair samples from two different people that I can’t tell apart.’ [The expert] made these statements of confidence, despite being unable to recall at trial the length or diameter of the one hair found on the bed.” *Id.*

⁹⁸ Office of Public Affairs, Justice Department Issues Draft Guidance Regarding Expert Testimony and Lab Reports in Forensic Science, June 3, 2016. These documents are known as the Uniform Language for Testimony and Reports. <https://www.justice.gov/opa/pr/justice-department-issues-draft-guidance-regarding>.

⁹⁹ Supporting Documentation for Department of Justice Proposed Uniform Language for Testimony and Reports for the Forensic Hair Examination Discipline at 4. www.justice.gov/dag/file/877741/download.

¹⁰⁰ WHITE HOUSE PCAST REPORT, *supra* note 33, at 13.

¹⁰¹ *Id.* DOJ’s supporting documents cited M.M. Houck & B. Budowle, *Correlation of Microscopic and Mitochondrial DNA Hair Comparisons*, 47 J. FORENSIC SCI. 964 (2002). This FBI study used mitochondrial DNA analysis to re-examine samples from previous FBI microscopic hair examination cases. The PCAST report did not accept that this study supported validity and reliability because the study showed that in 9 of 80 cases (11 percent) the microscopic examination found the hair indistinguishable but DNA analysis showed that the hairs came from different individuals.

guidelines were based on “foundational research” that PCAST questioned.

C. Firearms & Toolmark Identifications

Firearms identifications, popularly known as “ballistics,” is another long-established forensic discipline. It developed in the early part of the last century, and by the 1930s courts were admitting evidence based on this technique. Subsequent cases followed these precedents, admitting evidence of bullet, cartridge case, and shot shell identifications.¹⁰² Toolmark comparison, a related discipline, was also accepted during this period.¹⁰³ At the time *Daubert* was decided, the FBI’s position was clear: “Firearms identification is the Forensic Science discipline that identifies a bullet, cartridge case or other ammunition component as having been fired by a particular firearm *to the exclusion of all other firearms.*”¹⁰⁴ Yet, the examination, by means of a comparison microscope, is subjective and without a meaningful standard.

1. Post-Daubert Cases

The courts gave short shrift to the initial post-*Daubert* challenges to firearms and toolmark identifications.¹⁰⁵ In 2005, however, the legal landscape changed abruptly. In *United States v. Green*,¹⁰⁶ the district judge questioned the foundational basis of firearms identifications. The court wrote that the expert “declared that this match could be made ‘to the exclusion of every other firearm in the world.’ That conclusion, needless to say, is extraordinary, particularly given [his] data and methods.”¹⁰⁷ Moreover, the expert could not cite any reliable error rates and admitted that he relied mainly on his subjective judgment. In addition, “[t]here were no reference materials of any specificity, no national or even local database on which he relied. And although he relied on his past experience with these weapons, he had no notes or pictures memorializing his past

¹⁰² 1 GIANNELLI ET AL., *supra* note 15, § 14.06.

¹⁰³ *Id.* at § 14.12.

¹⁰⁴ FBI HANDBOOK OF FORENSIC SCIENCE 57 (rev. ed. 1994) (emphasis added).

¹⁰⁵ *See, e.g.*, *United States v. Hicks*, 389 F.3d 514, 526 (5th Cir. 2004) (stating that “the matching of spent shell casings to the weapon that fired them has been a recognized method of ballistics testing in this circuit for decades”); *United States v. Foster*, 300 F. Supp. 2d 375, 377 n.1 (D. Md. 2004) (“Ballistics evidence has been accepted in criminal cases for many years. . . . In the years since *Daubert*, numerous cases have confirmed the reliability of ballistics identification.”); *United States v. Santiago*, 199 F. Supp. 2d 101, 111 (S.D.N.Y. 2002) (“The Court has not found a single case in this Circuit that would suggest that the entire field of ballistics identification is unreliable.”).

¹⁰⁶ 405 F. Supp. 2d 104 (D. Mass. 2005).

¹⁰⁷ *Id.* at 107.

observations.”¹⁰⁸ In the end, the court restricted the expert’s testimony; he could only explain the ways in which the casings were similar but not that they came from a specific weapon “to the exclusion of every other firearm in the world.” In the court’s view, that conclusion “stretches well beyond [the expert’s] data and methodology.”¹⁰⁹ The court also cautioned: “The more courts admit this type of toolmark evidence without requiring documentation, proficiency testing, or evidence of reliability, the more sloppy practices will endure; we should require more.”¹¹⁰

A few weeks later, a different district judge in *United States v. Monteiro*¹¹¹ found that the technique “is largely a subjective determination [and] based on experience and expertise.”¹¹² Importantly, the court also concluded that the theory on which the expert relied was “tautological.” The Association of Firearm and Toolmark Examiners (AFTE), the leading organization of examiners, proposed the theory.¹¹³ Under this theory, the examiner may declare an identification if (1) there is “sufficient agreement” of marks between the crime scene and test bullets and (2) there is “sufficient agreement” when the examiner says there is.¹¹⁴ In short, the “sufficient agreement” threshold is “in the minds eye of the examiner and is based largely on training and experience.”¹¹⁵ The court would not admit the evidence unless the expert could better document the examination.

Together, *Green* and *Monteiro* should have served as a shot across the bow. But they did not; courts continued to admit the same evidence as before.¹¹⁶

¹⁰⁸ *Id.*

¹⁰⁹ *Id.* at 109.

¹¹⁰ *Id.*

¹¹¹ 407 F. Supp. 2d 351 (D. Mass. 2006).

¹¹² *Id.* at 355.

¹¹³ *See Theory of Identification, Association of Firearm and Toolmark Examiners*, 30 AFTE J. 86 (1998).

¹¹⁴ *See* Itiel E. Dror, *How Can Francis Bacon Help Forensic Science? The Four Idols of Human*, 50 JURIMETRICS 93, 104 (2009) (“The potential problem here is the nonscientific nature of the identification criteria. If the comparison of toolmarks enables conclusions about common origin when the unique surface contours of two toolmarks are in ‘sufficient agreement,’ what is the scientific definition and measurement of what constitutes such ‘sufficient agreement’? It seems that it is more in the eye of the beholder than strict scientific measures because it is determined without specific quantification and criteria.”).

¹¹⁵ *Monteiro*, 407 F. Supp. 2d at 370.

¹¹⁶ *See, e.g., United States v. Williams*, 506 F.3d 151, 161-162 (2d Cir. 2007) (upholding admissibility of firearms identification evidence-bullets and cartridge casings); *United States v. Natson*, 469 F. Supp. 2d 1253, 1261 (M.D. Ga. 2007) (“According to his testimony, these toolmarks were sufficiently similar to allow him to identify Defendant’s gun as the gun that fired the cartridge found at the crime scene. He opined that he held this opinion to a 100% degree of certainty The Court also finds [expert’s] opinions reliable and based upon a scientifically

2. NAS Ballistic Imaging Report (2008)

In 2008, the National Academy of Sciences published a report on computer imaging of bullets.¹¹⁷ Although firearms identification was not the primary focus of the investigation, a section of the report commented on the subject.¹¹⁸ After surveying the literature on uniqueness, reproducibility, and permanence of individual characteristics, the report noted that “[m]ost of these studies are limited in scale and have been conducted by firearms examiners (and examiners in training) in state and local law enforcement laboratories as adjuncts to their regular casework.”¹¹⁹ The report found that the “validity of the fundamental assumptions of uniqueness and reproducibility of firearms-related toolmarks has not yet been fully demonstrated.”¹²⁰ The report went on to caution:

Conclusions drawn in firearms identification should not be made to imply the presence of a firm statistical basis when none has been demonstrated. Specifically, . . . examiners tend to cast their assessments in bold absolutes, commonly asserting that a match can be made “to the exclusion of all other firearms in the world.” Such comments cloak an inherently subjective assessment of a match with an extreme probability statement that has no firm grounding and unrealistically implies an error rate of zero.¹²¹

Citing this report, the district court in *United States v. Glynn*¹²² ruled that the expert would only be permitted to testify that it was “more likely than not” that recovered bullets and cartridge cases came from a particular weapon.¹²³ The court also commented: “Based on the *Daubert* hearings . . . , the Court very

valid methodology. Evidence was presented at the hearing that the toolmark testing methodology he employed has been tested, has been subjected to peer review, has an ascertainable error rate, and is generally accepted in the scientific community.”)

¹¹⁷ NATIONAL RESEARCH COUNCIL, NATIONAL ACADEMY OF SCIENCES, BALLISTIC IMAGING (2008).

¹¹⁸ The committee was asked to assess the feasibility, accuracy, reliability, and technical capability of developing and using a national ballistic database as an aid to criminal investigations. It concluded: (1) “A national reference ballistic image database of all new and imported guns is not advisable at this time.” (2) The National Integrated Ballistics Information Network (NIBIN) “can and should be made more effective through operational and technological improvements.” *Id.*

¹¹⁹ *Id.* at 70.

¹²⁰ *Id.* at 81. The report also stated: “Additional general research on the uniqueness and reproducibility of firearms-related toolmarks would have to be done if the basic premises of firearms identification are to be put on a more solid scientific footing.” *Id.* at 82.

¹²¹ *Id.* at 82.

¹²² 578 F. Supp. 2d 567 (S.D.N.Y. 2008).

¹²³ *Id.* at 575.

quickly concluded that whatever else ballistics identification analysis could be called, it could not fairly be called ‘science,’¹²⁴ further noting that “[t]he problem is compounded by the tendency of ballistics experts . . . to make assertions that their matches are certain beyond all doubt, that the error rate of their methodology is ‘zero,’ and other such pretensions.”¹²⁵

3. *NAS Forensic Science Report (2009)*

As noted earlier, NAS issued its forensic report the following year in 2009. That report summarized the state of the research as follows:

Because not enough is known about the variabilities among individual tools and guns, we are not able to specify how many points of similarity are necessary for a given level of confidence in the result. Sufficient studies have not been done to understand the reliability and repeatability of the methods. . . . Individual patterns from manufacture or from wear might, in some cases, be distinctive enough to suggest one particular source, but additional studies should be performed to make the process of individualization more precise and repeatable.¹²⁶

In a different passage, the report — citing firearm and toolmark identifications — observed that “[m]uch forensic evidence . . . is introduced in criminal trials without any meaningful scientific validation, determination of error rates, or reliability testing to explain the limits of the discipline.”¹²⁷

AFTE rejected these findings out of hand, arguing that NAS “ignore[d] extensive research supporting the scientific underpinnings of the identification of firearm and toolmark evidence.”¹²⁸ The court in *United States v. Otero*¹²⁹ accepted the AFTE’s position, citing studies which it was ill-equipped to evaluate.¹³⁰ A subsequent review of the oft-cited studies by two scientists concluded:

Exaggerated and unfounded implications relating to rates of error inferred

¹²⁴ *Id.* at 570.

¹²⁵ *Id.* at 574.

¹²⁶ NAS FORENSIC REPORT, *supra* note 21, at 154.

¹²⁷ *Id.* at 107-08.

¹²⁸ *The Response of the Association of Firearms and Tool Mark Examiners to the February 2009 National Academy of Science Report “Strengthening the Forensic Science in the United States: A Path Forward,”* 41 AFTE J. 204, 206 (2009).

¹²⁹ 849 F. Supp. 2d 425, 437-38 (D.N.J. 2012) (“The Court’s analysis of the proposed testimony according to the *Daubert* factors leads it to conclude that [the] expert report and opinion are admissible under Rule 702.”).

¹³⁰ *See infra* notes 138-39 (PCAST report).

from even the best of existing experiments in the field of firearms/toolmarks, generally self-described as ‘validation studies’, typically result from statistical, metallurgical and/or psychological (cognitive) deficiencies in the design and conduct of the experiments, and frequently lead to unjustified inferential extrapolation to universal assumption for the practice domain.¹³¹

Other courts took an important, but still limited, step of restricting examiner testimony by precluding the expert from making gross overstatements such as declaring a match to the exclusion, either practical or absolute, of all other weapons.¹³² Similarly, some courts forbade experts from testifying that they hold their opinions to a “reasonable degree of scientific certitude.”¹³³ That term has long been required by courts in many jurisdictions for the admission of expert testimony. Incredibly, the phrase has no scientific meaning and the claim of certainty is unsupported by empirical research. Thus, it is grossly misleading. Indeed, the National Commission on Forensic Science rejected it.¹³⁴ Still other courts went off on a quixotic tangent, substituting the phrase “reasonable degree of ballistic” certitude.¹³⁵ Changing “scientific certainty” to “ballistic certainty” merely underscores the courts’ scientific incompetence.

¹³¹ Clifford H. Spiegelman & William A. Tobin, *Analysis of Experiments in Forensic Firearms/Toolmarks Practice Offered as Support for Low Rates of Practice Error and Claims of Inferential Certainty*, 13 LAW, PROB. & RISK 115, 115 (2013).

¹³² See, e.g., *United States v. Asburn*, 88 F. Supp. 3d 239, 249 (E.D.N.Y. 2015) (“Nor can [the expert] testify that a match he identified is to ‘the exclusion of all other firearms in the world,’ or that there is a ‘practical impossibility’ that any other gun could have fired the recovered materials.”); *United States v. Taylor*, 663 F. Supp. 2d 1170, 1180 (D.N.M. 2009) (“[The expert] also will not be allowed to testify that he can conclude that there is a match to the exclusion, either practical or absolute, of all other guns.”).

¹³³ See, e.g., *Asburn*, 88 F. Supp. 3d at 249 (“[T]he court joins in precluding this expert witness from testifying that he is ‘certain’ or ‘100%’ sure of his conclusions that certain items match.”); *United States v. Willock*, 696 F. Supp. 2d 536, 549 (D. Md. 2010) (“[The expert] shall state his opinions and conclusions without any characterization as to the degree of certainty with which he holds them.”); *People v. Robinson*, 2 N.E.3d 383, 402 (Ill. App. Ct. 2013) (“[T]he judicial decisions uniformly conclude toolmark and firearms identification is generally accepted and admissible at trial. Accordingly, we conclude the trial court did not err in ruling the testimony in this case was admissible . . . , particularly where the trial judge barred the witnesses from testifying their opinions were ‘within a reasonable degree of scientific certainty.’”).

¹³⁴ Nat’l Comm’n on Forensic Sci., Department of Justice, Views Document on Use of the Term “Reasonable Scientific Certainty” (adopted at NCFS Meeting #9 – March 22, 2016).

¹³⁵ *Taylor*, 663 F. Supp. 2d at 1180 (“He may only testify that, in his opinion, the bullet came from the suspect rifle to within a reasonable degree of certainty in the firearms examination field.”); *United States v. Cerna*, No. CR 08–0730 WHA, 2010 WL 3448528, at * 4 (N.D. Cal. 2010) (“a reasonable degree of certainty in the ballistics field”); *Commonwealth v. Pytou Heang*, 942 N.E.2d 927, 945 (Mass. 2011) (stating that “the expert may offer that opinion to a ‘reasonable degree of ballistic certainty’”).

However, even these modest limitations were rejected by other courts.¹³⁶ For example, in *United States v. Casey*,¹³⁷ the district court declined “to follow sister courts who have limited expert testimony based upon the 2008 and 2009 NAS reports and, instead, remains faithful to the long-standing tradition of allowing the unfettered testimony of qualified ballistics experts.”¹³⁸

4. *White House PCAST Report (2016)*

The 2016 White House PCAST report agreed with the NAS 2009 report’s characterization of the scientific research on firearms and toolmarks identification: “We find that many of these earlier studies were inappropriately designed to assess foundational validity and estimate reliability. Indeed, there is internal evidence among the studies themselves indicating that many previous studies underestimated the false positive rate by at least 100-fold.”¹³⁹ In addition, PCAST found only one of the post-2009 studies sufficiently rigorous. The Defense Department’s Forensic Science Center commissioned the study, which was conducted by an independent testing lab (the Ames Laboratory, a Department of Energy national laboratory affiliated with Iowa State University). In this study, “[t]he false-positive rate was estimated at 1 in 66, with a confidence bound indicating that the rate could be as high as 1 in 46.”¹⁴⁰ The study has not been published in a scientific journal. According to the PCAST report, more than one study is required and studies should be published in peer-reviewed scientific literature. Consequently, “the current evidence still falls short of the scientific criteria for foundational validity.”¹⁴¹

¹³⁶ See, e.g., *Fleming v. State*, 1 A.3d 572, 590 (Md. Ct. App. 2010) (“[N]otwithstanding the current debate on the issue, courts have consistently found the traditional method [of firearms identification] to be generally accepted within the scientific community, and to be reliable.”); *People v. Givens*, 912 N.Y.S.2d 855, 857 (Sup. Ct. 2010) (“This Court was unable to find any cases where firearms and toolmark identification was found to be unreliable or no longer scientifically acceptable.”).

¹³⁷ 928 F. Supp. 2d 397, 400 (D. Puerto Rico 2013).

¹³⁸ *Id.* at 400. See also *United States v. Sebborn*, No. 10 Cr. 87(SLT), 2012 WL 5989813 (E.D.N.Y. Nov. 30, 2012); *State v. Langlois*, 2 N.E.3d 936, ¶ 41 (Ohio Ct. App. 2014) (“Our conclusion on this issue finds support in the decisions of other appellate districts in Ohio, notwithstanding the recent criticisms in scientific reports and the limitations some federal courts have imposed on the testimony of firearms experts. These decisions hold that the methodology of comparatively analyzing and testing bullets and shell cases recovered from crime scenes is reliable.”); *State v. Jones*, 303 P.3d 1084, ¶ 75 (Wash. Ct. App. 2013) (expert testimony comparing bunter marks on the base of shell casings found at the crime scene to shell casings found in Jones’s home admissible under *Frye* standard).

¹³⁹ WHITE HOUSE PCAST REPORT, *supra* note 33, at 11.

¹⁴⁰ *Id.*

¹⁴¹ *Id.*

The AFTE quickly retorted, expressing their “disappointment in the PCAST’s choice to ignore the research that has been conducted” and claiming that “[d]ecades of validation and proficiency studies have demonstrated that firearm and toolmark identification is scientifically valid.”¹⁴² However, when PCAST later invited stakeholders to submit validation studies that it may have overlooked, no studies satisfying PCAST’s criteria were offered.¹⁴³

* * *

The lessons here are familiar. For years, an entrenched forensic discipline vigorously guarded its turf by rejecting the conclusions of the outside scientific community.¹⁴⁴ It published a journal which was “peer-reviewed” by other members of its discipline. The journal, which is advertised as “the Scientific Journal” of AFTE, was not generally available until 2016. The discipline claimed to be a “science” but did not hold itself to the normative standards of science. The AFTE “Theory of Identification” is “clearly not a scientific theory, which the National Academy of Sciences has defined as ‘a comprehensive explanation of some aspect of nature that is supported by a vast body of evidence. . . .’ More importantly, the stated method is circular.”¹⁴⁵ Only recently, after two NAS reports, have some courts begun to limit misleading testimony. Many have not. Thus, the courts’ competence to deal with flawed research remains extant.¹⁴⁶ The one bright spot came in *Williams v. United States*,¹⁴⁷ in which Judge Easterly wrote in a concurring opinion: “As matters currently stand, a certainty statement regarding toolmark pattern matching has the same probative value as the vision of a psychic: it reflects nothing more than the individual’s foundationless faith in what he believes to be true.”¹⁴⁸

D. Fingerprint Examinations

¹⁴² *Association of Firearm and Tool Mark Examiners Response to PCAST Report on Forensic Science*, 48 AFTE J. 195, 195 (2016).

¹⁴³ PRESIDENT’S COUNCIL OF ADVISORS ON SCIENCE AND TECHNOLOGY, AN ADDENDUM TO THE PCAST REPORT ON FORENSIC SCIENCE IN CRIMINAL COURTS, Jan. 6, 2017, at 7 (“Several respondents wrote to PCAST concerning firearms analysis. None cited additional appropriately designed black-box studies similar to the recent Ames Laboratory study.”).

¹⁴⁴ See William A. Tobin et al., *Absence of Statistical and Scientific Ethos: The Common Denominator in Deficient Forensic Practices*, 3 STATISTICS & PUBLIC POLICY (Dec. 16, 2016) (“[P]ractitioners remain intractable even after years of critical scholarly papers, ad hoc committees of the National Academy of Sciences (NAS), position statements from the U.S. Department of Justice . . .”).

¹⁴⁵ WHITE HOUSE PCAST REPORT, *supra* note 33, at 6.

¹⁴⁶ Tobin et al., *supra* note 144 (“the purported ‘validation studies’ typically proffered to courts are seriously flawed [and] have no external validity”).

¹⁴⁷ 130 A.3d 343 (D.C. 2016).

¹⁴⁸ *Id.* at 355 (concurring).

Before DNA analysis, fingerprint identification was the gold standard in forensics.¹⁴⁹ Like many other forensic disciplines, it gained judicial acceptance decades before *Daubert* was decided. *People v. Jennings*,¹⁵⁰ the first reported fingerprint case, was decided in 1911. In 1984, the FBI pronounced the technique “infallible” in its official publication, which also referred to the technique as a “science.”¹⁵¹ Nevertheless, it is a subjective technique without an objective standard and typically involves partial prints with inevitable distortions.

1. *Post-Daubert Cases*

After *Daubert*, challenges to fingerprint comparison testimony were decidedly unsuccessful.¹⁵² One infamous case, *United States v. Havvard*,¹⁵³ illustrates the judiciary’s lack of rigor in applying *Daubert*. Not only did the district court uphold the admissibility of fingerprint testimony, it described the technique as “the very archetype of reliable expert testimony under [the *Daubert/Kumho*] standards.”¹⁵⁴ According to the court, latent print identification had been “tested” for nearly 100 years in adversarial proceedings with the highest possible stakes — liberty and sometimes life. Yet, *Daubert* required scientific, not “adversarial,” testing.¹⁵⁵ Next, in citing “peer review,” the court noted that a second fingerprint examiner also compared the prints: “In fact, peer review is the standard operating procedure among latent print examiners.”¹⁵⁶ This statement reveals a fundamental misunderstanding of “peer review” as used in *Daubert*. In that case, peer review meant refereed scientific journals in which validation research is published. An amici brief submitted in *Daubert* by the *New England Journal of Medicine* and other scientific publications explained that peer review’s “role is to promote the publication of well-conceived articles so that the most

¹⁴⁹ See Joseph L. Peterson & Anna S. Leggett, *The Evolution of Forensic Science: Progress Amid the Pitfalls*, 36 STETSON L. REV. 621, 654 (2007) (“The scientific integrity and reliability of DNA testing have helped DNA replace fingerprinting and made DNA evidence the new ‘gold standard’ of forensic evidence.”).

¹⁵⁰ 96 N.E. 1077 (Ill. 1911). See generally 1 GIANNELLI ET AL., *supra* note 15, ch. 18 (discussing the scientific and legal issues associated with fingerprint identification).

¹⁵¹ FEDERAL BUREAU OF INVESTIGATION, THE SCIENCE OF FINGERPRINTS iv (1984).

¹⁵² See, e.g., *United States v. Collins*, 340 F.3d 672, 682 (8th Cir. 2003) (“Fingerprint evidence and analysis is generally accepted.”); *United States v. Hernandez*, 299 F.3d 984, 991 (8th Cir. 2002); *United States v. Martinez-Cintrón*, 136 F. Supp. 2d 17, 20 (D. Puerto Rico 2001).

¹⁵³ 117 F. Supp. 2d 848 (S.D. Ind. 2000), *aff’d*, 260 F.3d 597 (7th Cir. 2001).

¹⁵⁴ *Id.* at 855.

¹⁵⁵ See Sandy L. Zabell, *Fingerprint Evidence*, 13 J.L. & POL’Y 143, 170 (2005) (The “argument that no latent print has ever been found to match the rolled print of a different person is . . . misleading because no systematic search for such pairs on the entire databank of millions of fingerprints has ever been performed.”).

¹⁵⁶ *Havvard*, 117 F. Supp. 2d at 854.

important review, the consideration of the reported results by the scientific community, may occur after publication.”¹⁵⁷

Moreover, the court accepted the prosecution expert’s astounding claim that the “error rate for the method is zero.”¹⁵⁸ Experts argued that, while individual examiners may make mistakes, the method itself is perfect. However, the dichotomy between “methodological” and “human” error rates in this context is “practically meaningless”¹⁵⁹ because the examiner is the method.¹⁶⁰ Finally, the court turned *Daubert* on its head by requiring the defendant to prove the evidence was unreliable, a distortion that would be employed in later cases.¹⁶¹

Then, *United States v. Llera Plaza*¹⁶² “sent shock waves through the community of fingerprint analysts.”¹⁶³ In that 2002 case, Judge Pollak ruled that fingerprint experts would not be permitted to testify that two sets of prints “matched” — that is, a positive identification to the exclusion of all other persons. This was apparently the first time in over 90 years that such a decision had been rendered.¹⁶⁴ On rehearing, however, Judge Pollak reversed himself,¹⁶⁵

¹⁵⁷ Brief of the New England Journal of Medicine, Journal of the American Medical Association, and Annals of Internal Medicine as Amici Curiae in Support of Respondent, *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579 (1993), at 3.

¹⁵⁸ *Havvard*, 117 F. Supp. 2d at 854.

¹⁵⁹ Jennifer L. Mnookin, *Fingerprint Evidence in an Age of DNA Profiling*, 67 BROOK. L. REV. 13, 60 (2001). Professor Mnookin goes on to provide this analogy: “The same argument could be made of eyewitness testimony, a notoriously unreliable form of evidence. People are all distinct from one another in observable ways; therefore the theoretical error rate of eyewitness identification is zero, though in practice observers may frequently makes errors.” *Id.* See also Simon A. Cole, *More Than Zero: Accounting for Error in Latent Fingerprint Identification*, 95 J. CRIM. L. & CRIMINOLOGY 985, 1040 (2005) (stating that “in fingerprint practice the concept is vacuous”).

¹⁶⁰ See Zabell, *supra* note 155, at 172 (“But, given its unavoidable subjective component, in latent print examination people *are* the process.”).

¹⁶¹ See Michael J. Saks, *The Legal and Scientific Evaluation of Forensic Science (Especially Fingerprint Expert Testimony)*, 33 SETON HALL L. REV. 1167, 1173-76 (2003) (discussing the reversal of the burden of persuasion as one of several judicial responses employed to avoid confronting the lack of empirical testing).

¹⁶² 179 F. Supp. 2d 492 (E.D. Pa.), *vacated, mot. granted on recons.*, 188 F. Supp. 2d 549 (E.D. Pa. 2002).

¹⁶³ See D.H. Kaye, *The Nonscience of Fingerprinting: United States v. Llera-Plaza*, 21 QUINNIPIAC L. REV. 1073, 1073 (2003).

¹⁶⁴ As Professor Mnookin has noted, however, “fingerprints were accepted as an evidentiary tool without a great deal of scrutiny or skepticism.” Mnookin, *supra* note 159, at 17. She elaborated: “Even if no two people had identical *sets* of fingerprints, this did not establish that no two people could have a *single* identical print, much less an identical *part* of a print. These are necessarily matters of probability, but neither the court in *Jennings* nor subsequent judges ever required that fingerprinting identification be placed on a secure statistical foundation.” *Id.* at 19.

¹⁶⁵ 188 F. Supp. 2d 549, 572 (E.D. Pa. 2002).

and later cases continued to uphold the admissibility of fingerprint evidence.¹⁶⁶ Nevertheless, the case captured the attention of the media with news reports,¹⁶⁷ mainstream publications,¹⁶⁸ scientific journals,¹⁶⁹ and television shows giving it substantial coverage.¹⁷⁰ A spate of legal articles followed,¹⁷¹ with some commentators believing that *Llera Plaza I* was more faithful to *Daubert* than *Llera Plaza II*.¹⁷² In response, the FBI adopted a “circle the wagons” attitude, fiercely defending the technique. The head of the FBI fingerprint section told *60 Minutes* that the error rate was “zero”, examiners only testify to “hundred percent certainty,” and the FBI had won “forty-one out of forty-one” legal challenges to fingerprint evidence.¹⁷³

The appellate opinion most faithful to *Daubert* appeared in *United States v. Crisp*¹⁷⁴ — unfortunately in dissent. The majority opinion upheld the admissibility of fingerprint evidence by shifting the burden of proof to the defendant and by grandfathering the technique.¹⁷⁵ In dissent, Judge Michael

¹⁶⁶ See, e.g., *United States v. Abreu*, 406 F.3d 1304, 1307 (11th Cir. 2005) (“We agree with the decisions of our sister circuits and hold that the fingerprint evidence admitted in this case satisfied *Daubert*.”); *United States v. Janis*, 387 F.3d 682, 690 (8th Cir. 2004) (finding fingerprint evidence reliable); *United States v. Mitchell*, 365 F.3d 215 (3d Cir. 2004); *United States v. Crisp*, 324 F.3d 261 (4th Cir. 2003); *United States v. Sullivan*, 246 F. Supp. 2d 700, 704 (E.D. Ky. 2003).

¹⁶⁷ E.g., Joann Loviglio, *Trial Judge Reaffirms Fingerprint Usability; Hearing Shows Him Science Involved*, SAN ANTONIO EXPRESS-NEWS, Mar. 14, 2002; Andy Newman, *Judge Who Ruled Out Matching Fingerprints Changes his Mind*, N.Y. TIMES, Mar. 14, 2002; Richard Willing, *Judge Challenges Fingerprint Identification*, USA TODAY, Jan. 10, 2002.

¹⁶⁸ E.g., Michael Specter, *Do Fingerprints Lie? The Gold Standard of Forensic Science is Now Being Challenged*, 78 THE NEW YORKER 96 (May 27, 2002) (discussing case including interview with judge).

¹⁶⁹ See David L. Faigman, *Is Science Different for Lawyers?*, 297 SCIENCE 339, 340 (2002).

¹⁷⁰ *60 Minutes: Fingerprints* (CBS television broadcast Jan. 5, 2003).

¹⁷¹ See, e.g., Simon A. Cole, *Grandfathering Evidence: Fingerprint Admissibility Rulings from Jennings to Llera Plaza and Back Again*, 41 AM. CRIM. L. REV. 1189 (2004); Robert Epstein, *Fingerprints Meet Daubert: The Myth of Fingerprint “Science” Is Revealed*, 75 S. CAL. L. REV. 605 (2002); Kristin Romandetti, Note, *Recognizing and Responding to a Problem with the Admissibility of Fingerprint Evidence Under Daubert*, 45 JURIMETRICS 41 (2004).

¹⁷² E.g., Jennifer L. Mnookin, *Fingerprints: Not a Gold Standard*, 20 ISSUES IN SCI. & TECH. 47 (2003) (“Judge Pollak’s first opinion [restricting latent fingerprint individualization testimony] was the better one.”); Recent Case, 115 HARV. L. REV. 2349, 2352 (2002) (“Fingerprint expert testimony does not survive application of the *Daubert* factors . . .”).

¹⁷³ *60 Minutes: Fingerprints* (CBS television broadcast Jan. 5, 2003).

¹⁷⁴ 324 F.3d 261 (4th Cir. 2003).

¹⁷⁵ *Id.* at 269 (“Put simply, Crisp has provided us no reason today to believe that this general acceptance of the principles underlying fingerprint identification has, for decades, been misplaced. Accordingly, the district court was well within its discretion in accepting at face value the consensus of the expert and judicial communities that the fingerprint identification

conscientiously applied the *Daubert* factors. First, he noted that the “government did not offer any record of testing on the reliability of fingerprint identification. . . . [T]here have not been any studies to establish how likely it is that partial prints taken from a crime scene will be a match for only one set of fingerprints in the world.”¹⁷⁶ Second, as for peer review, “[a]gain, the government offered no evidence on this factor at trial. Fingerprint examiners, . . . have their own professional publications. . . . But unlike typical scientific journals, the fingerprint publications do not run articles that include or prompt critique or reanalysis by other scientists. Indeed, few of the articles address the principles of fingerprint analysis and identification at all”¹⁷⁷ Third, “an error rate must be demonstrated by reliable scientific studies, not by assumption.”¹⁷⁸ Fourth, “the government did not establish that there are objective standards in the fingerprint examination field to guide examiners in making their comparisons.”¹⁷⁹ Fifth, while acknowledging general acceptance in the fingerprint community, the judge remarked that “[n]othing in the record in this case shows that the fingerprint examination community has challenged itself sufficiently or has been challenged in any real sense by outside scientists.”¹⁸⁰

In conclusion, the judge wrote: “The government has had ten years to comply with *Daubert*. It should not be given a pass in this case.”¹⁸¹

2. Madrid Train Bombing

Llera Plaza was soon eclipsed by a more sensational event — the FBI’s misidentification of Brandon Mayfield as the source of the crime scene prints in the terrorist train bombing in Madrid on March 11, 2004.¹⁸² More than any other event, the Mayfield affair exposed the myth of fingerprint infallibility. This

technique is reliable.”).

¹⁷⁶ *Id.* at 273-74 (Michael, J., dissenting).

¹⁷⁷ *Id.* at 274.

¹⁷⁸ *Id.* The judge added: “In a 1995 test conducted by a commercial testing service, less than half of the fingerprint examiners were able to identify correctly all of the matches and eliminate the non-matches. On a similar test in 1998, less than sixty percent of the examiners were able to make all identifications and eliminations. . . . An error rate that runs remarkably close to chance can hardly be viewed as acceptable under *Daubert*.” *Id.* at 275.

¹⁷⁹ *Id.* at 276.

¹⁸⁰ *Id.*

¹⁸¹ *Id.* at 272.

¹⁸² See Sara Kershaw, *Spain and U.S. at Odds on Mistaken Terror Arrest*, N.Y. TIMES, Jun. 5, 2004, at A1 (Spanish authorities cleared Brandon Mayfield and matched the fingerprints to an Algerian national); Flynn McRoberts & Maurice Possley, *Report Blasts FBI Lab: Peer Pressure Led to False ID of Madrid Fingerprint*, CHI. TRIB., Nov. 14, 2004, at 1.

debacle resulted in investigations by the FBI¹⁸³ and the Inspector General (IG) of the Department of Justice.¹⁸⁴ One of the more troubling aspects of these reports dealt with the culture in the laboratory. The FBI internal investigation found that “[t]o disagree was not an expected response,”¹⁸⁵ and the IG reported that “FBI examiners did not attempt to determine the basis of the [Spanish National Police’s] doubts before reiterating that they were ‘absolutely confident’ in the identification on April 15, a full week before the FBI Laboratory met with the SNP.”¹⁸⁶

In addition to highlighting the lack of foundational research, these events raised a host of other issues, including (1) the role of cognitive bias in subjective techniques,¹⁸⁷ (2) the lack of well-defined standards,¹⁸⁸ (3) the failure to administer rigorous proficiency tests,¹⁸⁹ (4) the manipulation of research,¹⁹⁰ and (5) other

¹⁸³ See Robert B. Stacey, *A Report on the Erroneous Fingerprint Individualization in the Madrid Train Bombing Case*, 54 J. FORENSIC IDENTIFICATION 707 (2004).

¹⁸⁴ See OFFICE OF THE INSPECTOR GENERAL, U.S. DEP’T OF JUSTICE, A REVIEW OF THE FBI’S HANDLING OF THE BRANDON MAYFIELD CASE, UNCLASSIFIED EXECUTIVE SUMMARY 7 (2006) (“Having found as many as 10 points of unusual similarity, the FBI examiners began to ‘find’ additional features in [the print] that were not really there, but rather were suggested to the examiners by features in the Mayfield prints.”).

¹⁸⁵ Stacey, *supra* note 183, at 713.

¹⁸⁶ OFFICE OF THE INSPECTOR GENERAL, *supra* note 184, at 10.

¹⁸⁷ See Itiel E. Dror et al., *Contextual Information Renders Experts Vulnerable to Making Erroneous Identifications*, 156 FORENSIC SCI. INT’L 74 (2006) (reporting an experiment that showed fingerprint examiners changed their opinions when provided with irrelevant information); Elizabeth F. Loftus & Simon A. Cole, Letter, *Contaminated Evidence*, 304 SCIENCE 959 (May 14, 2004) (“[F]orensic scientists remain stubbornly unwilling to confront and control the problem of bias, insisting that it can be overcome through sheer force of will and good intentions.”); Stacey, *supra* note 185, at 713 (“confirmation bias”).

¹⁸⁸ Examiners follow a procedure known as Analysis, Comparison, Evaluation, and Verification (ACE-V). See Zabell, *supra* note 155, at 178 (“ACE-V is an acronym, not a methodology. It is merely the common sense description of what anyone would do if they were examining a latent and a candidate source print.”).

¹⁸⁹ See *Crisp*, 324 F.3d at 274 (4th Cir. 2003) (Michael, J., dissenting) (“Proficiency testing is typically based on a study of prints that are far superior to those usually retrieved from a crime scene.”); *Llera Plaza*, 188 F. Supp. 2d at 565 (noting that “the FBI examiners got very high proficiency grades, but the tests they took did not. . . . [O]n the present record I conclude that the proficiency tests are less demanding than they should be.”); Jennifer L. Mnookin, Editorial, *A Blow to the Credibility of Fingerprint Evidence*, BOSTON GLOBE, Feb. 2, 2004 (“There are no systematic proficiency tests to evaluate examiners’ skill. Those tests that exist are not routinely used and are substandard.”).

¹⁹⁰ See Donald Kennedy, Editorial, *Forensic Science: Oxymoron?*, 302 SCIENCE 1625 (2003) (discussing the cancellation of a National Academies project designed to examine various forensic science techniques, including fingerprinting, because the Departments of Justice and Defense insisted on a right of review that the Academy had refused to other grant sponsors); *United States v. Mitchell*, 365 F.3d 215, 238 (3d Cir. 2004) (“We are deeply discomfited by Mitchell’s contention — supported by Dr. Rau’s account of events, though contradicted by other

instances of misidentifications.¹⁹¹ The FBI did not undertake a serious review of fingerprints until it was compelled to address the issue due to the negative publicity surrounding the Mayfield misidentification. Even then, however, the FBI still characterized the technique as “scientific.”¹⁹²

The scientific community continued to note the lack of research,¹⁹³ — and the courts continued to ignore this fact.¹⁹⁴ Indeed, in *United States v. Baines*,¹⁹⁵ decided in 2009, the head of the FBI fingerprint section testified: “As to these ‘false positives’ . . . the FBI had ‘made, on average, about one erroneous identification every 11 years.’ The total number of identifications made has been about one million per year, . . . so that the known actual error rate was about one per eleven million identifications.”¹⁹⁶ Problematically, he merely assumed that all the other identifications were correct, thus disqualifying his analysis. Perhaps the

witnesses — that a conspiracy within the Department of Justice intentionally delayed the release of the solicitation until after Mitchell’s jury reached a verdict. Dr. Rau’s story, if true, would be a damning indictment of the ethics of those involved.”). See generally Paul C. Giannelli, *Daubert and Forensic Science: The Pitfalls of Law Enforcement Control of Scientific Research*, 2011 U. ILLINOIS L. REV. 53 (discussing the manipulation of forensic science research, including fingerprint research).

¹⁹¹ See Simon A. Cole, *More Than Zero: Accounting for Error in Latent Fingerprint Identification*, 95 J. CRIM. L. & CRIMINOLOGY 985 (2005) (collecting 23 cases involving mistakes). The misidentification cases include some that involved (1) verification by one or more other examiners, (2) examiners certified by the International Association of Identification, (3) procedures using a sixteen-point standard, and (4) defense experts who corroborated misidentifications made by prosecution experts. *Id.* at 985; *Reasonable Doubt: Can We Trust Crime Labs?*, CNN PRESENTS, Jan. 9, 2005 (discussing the misidentification of Ricki Jackson, who spent two years in prison).

¹⁹² See Bruce Budowle et al., *Review of the Scientific Basis for Friction Ridge Comparisons as a Means of Identification: Committee Findings and Recommendations*, 8 FORENSIC SCI. COMM. (Jan. 2006).

¹⁹³ See Donald Kennedy & Richard A. Merrill, *Assessing Forensic Science*, 20 ISSUES IN SCI. & TECH. 33 (Fall 2003) (“The increased use of DNA analysis, which has undergone extensive validation, has thrown into relief the less firmly credentialed status of other forensic science identification techniques (fingerprints, fiber analysis, hair analysis, ballistics, bite marks, and tool marks). These have not undergone the type of extensive testing and verification that is the hallmark of science elsewhere.”); Zabell, *supra* note 155, at 164 (“Although there is a substantial literature on the uniqueness of fingerprints, it is surprising how little true scientific support for the proposition exists.”).

¹⁹⁴ See, e.g., *United States v. Pena*, 586 F.3d 105, 110 (1st Cir. 2009) (“The district court did not abuse its discretion. Numerous courts have found expert testimony on fingerprint identification based on the ACE-V method to be sufficiently reliable under *Daubert*.”); *United States v. Abreu*, 406 F.3d 1304, 1307 (11th Cir. 2005) (“[T]he fingerprint evidence admitted in this case satisfied *Daubert*.”); *United States v. Janis*, 387 F.3d 682, 690 (8th Cir. 2004) (“We conclude the district court did not err in admitting the fingerprint expert’s testimony.”); *United States v. Mitchell*, 365 F.3d 215, 241 (3d Cir. 2004).

¹⁹⁵ 573 F.3d 979 (10th Cir. 2009).

¹⁹⁶ *Id.* at 984.

most troubling aspect of this testimony was the lack of self-awareness for a person who claimed to be a scientist.¹⁹⁷

3. *NAS Forensic Science Report (2009)*

Fingerprint examiners follow a procedure known as Analysis, Comparison, Evaluation, and Verification (ACE-V). The 2009 NAS report observed that since “the ACE-V method does not specify particular measurements or a standard test protocol, . . . examiners must make subjective assessments throughout.”¹⁹⁸ Thus, the ACE-V method is too “broadly stated” to “qualify as a validated method for this type of analysis.”¹⁹⁹ The report added that “[t]he latent print community in the United States has eschewed numerical scores and corresponding thresholds” and consequently relies “on primarily subjective criteria” in making the ultimate attribution decision.²⁰⁰ In making the decision, the examiner must draw on his or her personal experience to evaluate such factors as “inevitable variations in pressure,” but to date those factors have not been “characterized, quantified, or compared.”²⁰¹ In addition, the report gave short shift to the zero-error-rate argument, finding that “claims that these analyses have zero error rates are not scientifically plausible.”²⁰² In conclusion, the report outlined an agenda for the research it considered necessary “[t]o properly underpin the process of friction ridge identification.”²⁰³

Several studies were published after the NAS report.²⁰⁴ The most important was a FBI study published in 2011,²⁰⁵ which is discussed below.

¹⁹⁷ See WHITE HOUSE PCAST REPORT, *supra* note 33, at 53 (“The fallacy is obvious: the expert simply assumed without evidence that every error in casework had come to light.”).

¹⁹⁸ NAS FORENSIC REPORT, *supra* note 21, at 139.

¹⁹⁹ *Id.* at 142.

²⁰⁰ *Id.* at 141.

²⁰¹ *Id.* at 144. Moreover, examiners lack population frequency data to quantify how rare or common a particular type of fingerprint characteristic is. *Id.* at 144.

²⁰² *Id.* at 142. See also *id.* at 143 (“Some in the latent print community argue that the method itself, if followed correctly . . . has a zero error rate. Clearly, this assertion is unrealistic The method, and the performance of those who use it, are inextricably linked, and both involve multiple sources of error (e.g., errors in executing the process steps, as well as errors in human judgment.”).

²⁰³ *Id.* at 144.

²⁰⁴ See WHITE HOUSE PCAST REPORT, *supra* note 33, at 91-95.

²⁰⁵ B.T. Ulery et al., *Accuracy and Reliability of Forensic Latent Fingerprint Decisions*, 108 PROC. NAT’L ACAD. SCI. 7733 (2011) (“To attempt to ensure that the non-mated pairs were representative of the type of matches that might arise when police identify a suspect by searching fingerprint databases, the known prints were selected by searching the latent prints against the 58 million fingerprints in the [Automated Fingerprint Identification System] database

4. *White House PCAST Report (2016)*

According to the White House PCAST report, “latent fingerprint analysis is a foundationally valid subjective methodology”²⁰⁶ and the FBI “significantly advanced the field” by conducting the black-box study. Nevertheless, the false positive rate

is substantial and is likely to be higher than expected by many jurors based on longstanding claims about the infallibility of fingerprint analysis. The false-positive rate could be as high as 1 error in 306 cases based on the FBI study and 1 error in 18 cases based on a study by another crime laboratory. In reporting results of [a] latent-fingerprint examination, it is important to state the false-positive rates based on properly designed validation studies[.]²⁰⁷

Moreover, “testimony asserting any specific level of increased accuracy (beyond that measured in the studies) due to blind independent verification would be scientifically inappropriate, as speculation unsupported by empirical evidence.”²⁰⁸

* * *

Despite the ruckus created by *Llera Plaza* and the Mayfield fiasco, examiner testimony has remained unchanged. Testimony such as “zero error rates,” “matches to the exclusion of all other fingerprints,” and “100 percent certainty” — which had been used for decades — has continued, while the fingerprint community remain oblivious that such statements were scientifically implausible.

On a positive note, the Mayfield incident did trigger the FBI’s black box study, which was a significant achievement. Still, this study was released 100 years after the courts first admitted fingerprint evidence.²⁰⁹ The White House PCAST report found it “distressing” that properly constructed validation studies had only been conducted recently and only one study had been published in a

and selecting one of the closest matching hits.”).

²⁰⁶ WHITE HOUSE PCAST REPORT, *supra* note 33, at 9-10.

²⁰⁷ *Id.* at 10.

²⁰⁸ *Id.* at 99.

²⁰⁹ *See supra* text accompanying note 150.

peer-reviewed journal.²¹⁰ *Daubert* had little effect.²¹¹ It took a serendipitous event — the Madrid train bombing — to compel the research.

E. Comparative Bullet Lead Analysis

For over thirty years, FBI experts testified about Comparative Bullet Lead Analysis (CBLA), a technique that was first used in the investigation into President Kennedy’s assassination.²¹² CBLA compares trace chemicals found in bullets at crime scenes with ammunition found in the possession of a suspect. This technique was used when firearms (“ballistics”) identification could not be employed. FBI experts used various analytical techniques (first, neutron activation analysis (NAA), and then inductively coupled plasma-atomic emission spectrometry (ICP-AES)) to determine the concentrations of seven elements — arsenic, antimony, tin, copper, bismuth, silver, and cadmium — in the bullet lead alloy of both the crime-scene and suspect’s bullets. Statistical tests were then used to compare the elements in each bullet and determine whether the fragments and suspect’s bullets were “analytically indistinguishable” for each of the elemental concentration means. Exactly what the phrase “analytically indistinguishable” meant was the central issue — i.e., did such a finding mean that the bullet fragments came from a small or large universe? Obviously, the probative value of the test results would differ if only a hundred bullets had the same chemical composition as opposed to several million bullets.

The published cases revealed disparate and often inconsistent interpretive conclusions provided by FBI experts. In some, experts testified only that two exhibits were “analytically indistinguishable.”²¹³ In other cases, examiners concluded that samples *could have* come from the same “source” or “batch.”²¹⁴ In still others, they stated that the samples *came* from the same source.²¹⁵ The testimony in numerous cases went much further and referred to a “box” of

²¹⁰ WHITE HOUSE PCAST REPORT, *supra* note 33, at 95.

²¹¹ Some courts did place limitations on the testimony. *See, e.g., Mitchell*, 365 F.3d at 245-46 (“Testimony at the *Daubert* hearing indicated that some latent fingerprint examiners insist that there is no error rate associated with their activities or that the examination process is irreducibly subjective. This would be out-of-place under Rule 702.”); *Commonwealth v. Gambora*, 933 N.E.2d 50, 61 n.22 (Mass. 2010) (“opinions expressing absolute certainty about, or the infallibility of, an ‘individualization’ of a print should be avoided”).

²¹² *See generally* Erik Randich & Patrick M. Grant, *Proper Assessment of the JFK Assassination Bullet Lead Evidence from Metallurgical and Statistical Perspectives*, 51 J. FORENSIC SCI. 717 (2006) (discussing the original analysis of the bullet fragments).

²¹³ *See Wilkerson v. State*, 776 A.2d 685, 689 (Md. 2001).

²¹⁴ *See State v. Krummacher*, 523 P.2d 1009, 1012-13 (Or. 1974).

²¹⁵ *See United States v. Davis*, 103 F.3d 660, 673-74 (8th Cir. 1996); *People v. Lane*, 628 N.E.2d 682, 689-90 (Ill. App. Ct. 1993).

ammunition (usually 50 loaded cartridges, sometimes 20). For example, two specimens:

- (1) Could have come from the same box,²¹⁶
- (2) Could have come from the same box or a box manufactured on the same day,²¹⁷
- (3) Were consistent with their having come from the same box of ammunition,²¹⁸
- (4) Probably came from the same box,²¹⁹ or
- (5) Must have come from the same box or from another box that would have been made by the same company on the same day.²²⁰

Several other statements that differ appear in the published opinions. An early case reported that the specimens “had come from the same batch of ammunition: they had been made by the same manufacturer on the same day and at the *same hour*.”²²¹ One case reports the expert’s conclusion with a statistic.²²² In another case, the expert used the expressions “such a finding is rare”²²³ and “a very rare finding.”²²⁴ In still another case, the expert “opined that the same company produced the bullets at the same time, using the same lead source. Based upon Department of Justice records, she opined that an overseas company called PMC produced the bullets around 1982.”²²⁵

²¹⁶ See *State v. Strain*, 885 P.2d 810, 817 (Utah Ct. App. 1994); *State v. Jones*, 425 N.E.2d 128, 131 (Ind. 1981).

²¹⁷ See *State v. Grube*, 883 P.2d 1069, 1078 (Idaho 1994); *People v. Johnson*, 499 N.E.2d 1355, 1366 (Ill. 1986).

²¹⁸ See *State v. Reynolds*, 297 S.E.2d 532, 534 (N.C. 1982).

²¹⁹ See *Bryan v. State*, 935 P.2d 338, 360 (Okla. Crim. App. 1997).

²²⁰ See *United States v. Davis*, 103 F.3d 660, 666-67 (8th Cir. 1996) (“An expert testified that such a finding is rare and that the bullets must have come from the same box or from another box that would have been made by the same company on the same day.”); *Commonwealth v. Daye*, 587 N.E.2d 194, 207 (Mass. 1992); *State v. King*, 546 S.E.2d 575, 584 (N.C. 2001) (The expert “opined that, based on her lead analysis, the bullets she examined either came from the same box of cartridges or came from different boxes of the same caliber, manufactured at the same time.”).

²²¹ *Brown v. State*, 601 P.2d 221, 224 (Alaska 1979) (emphasis added).

²²² *State v. Earhart*, 823 S.W.2d 607, 614 (Tex. Crim. App. 1991).

²²³ *United States v. Davis*, 103 F.3d 660, 666 (8th Cir. 1996).

²²⁴ *Id.* at 667.

²²⁵ *People v. Villarta*, No. H021354, 2002 WL 66887 (Cal. Ct. App. Jan. 17, 2002) (murder). In later years, the testimony became more limited. A 2002 FBI publication states the conclusion as follows: “Therefore, they *likely* originated from the same manufacturer’s source (melt) of lead.” Charles A. Peters, *The Basis for Compositional Bullet Lead Comparisons*, 4 FORENSIC SCI. COMM. No. 3, at 5 (July 2002) (emphasis added). Testimony to the same effect has also been proffered. Transcript of Record at 6, *Commonwealth v. Wilcox*, No. 00CR2727 (Ky. Cir. Ct. Jefferson County Feb. 28, 2002) (trial testimony of Charles Peters, FBI examiner): “Well,

1. *NAS Bullet Lead Report (2004)*

The technique was not seriously challenged until a retired FBI examiner, William Tobin, began questioning the procedure in scientific and legal journals²²⁶ and in court testimony as well.²²⁷ As a result, the FBI asked the National Academy of Sciences to review the technique. The 2004 NAS report undercut the FBI testimony: “The available data do not support any statement that a crime bullet came from a particular box of ammunition. In particular, references to ‘boxes’ of ammunition in any form should be avoided as misleading under Federal Rule of Evidence 403.”²²⁸ Perhaps the most disturbing case is *State v. Earhart*,²²⁹ a capital murder case in which the CBLA evidence apparently played a significant role.²³⁰ The transcript contains the following expert testimony: “We can — from my 21 years experience of doing bullet lead analysis and doing research on boxes of ammunition down though the years I can determine if bullets came from the same box of ammunition”²³¹ However, the NAS report found

bullets that are analytically indistinguishable likely come from the same molten lead sources of lead, uh, as opposed to bullets that have different composition come from different, uh, melts of lead.”

²²⁶ See Edward J. Imwinkelried & William A. Tobin, *Comparative Bullet Lead Analysis (CBLA) Evidence: Valid Inference or Ipse Dixit?*, 28 OKLA. CITY U. L. REV. 43 (2003); Erik Randich et al., *A Metallurgical Review of the Interpretation of Bullet Lead Compositional Analysis*, 127 FORENSIC SCI. INT’L 174 (2002) (Tobin was a coauthor); William A. Tobin & Wayne Duerfeldt, *How Probative is Comparative Bullet Lead Analysis?*, 17 CRIM. JUSTICE 26 (Fall 2002).

²²⁷ E.g., *Ragland v. Commonwealth*, 191 S.W.3d 569, 577 (Ky. 2006); *Clemons v. State*, 896 A.2d 1059, 1070, 1068 (Md. 2006); *State v. Behn*, 868 A.2d 329, 339-40 (N.J. Super. Ct. 2005) (Tobin’s affidavit submitted).

²²⁸ NATIONAL RESEARCH COUNCIL, NATIONAL ACADEMY OF SCIENCES, FORENSIC ANALYSIS: WEIGHING BULLET LEAD EVIDENCE 6 (2004). The author served on the NAS Committee.

²²⁹ 823 S.W.2d 607, 614 (Tex. Crim. App. 1991) (“[The expert] concluded that the likelihood that two .22 caliber bullets came from the same batch, based on *all* the .22 bullets made in one year, is approximately .000025 percent, ‘give or take a zero.’ He subsequently acknowledged, however, that the numbers which he used to reach the .000025 percent statistic failed to take into account that there are different types of .22 caliber bullets made each year – .22, .22 long, and .22 long rifle. [The expert] ultimately testified that there could be several hundred thousand bullets per batch, but with some variation in the elemental composition within the batch.”) (emphasis added).

²³⁰ See *Earhart v. Johnson*, 132 F.3d 1062, 1067 (5th Cir. 1998) (denying habeas relief, the court noted: “Given the significant role the bullet evidence played in the prosecution’s case, we shall therefore assume Earhart could have made a sufficient threshold showing that he was entitled to a defense expert under Texas law.”).

²³¹ Transcript of Record at 5248-49, *State v. Earhart*, No. 4064, Dist. Ct. Lee County, 21st Judicial Dist., Texas (testimony of John Riley). See also *id.* at 5258 (“Well, bullets that are — that have analytically indistinguishable compositions or compositions that are generally similar typically are found within the same box of ammunition and that is the case that we have

that the amount of bullets that can be produced from a melt “can range from the equivalent of as few as 12,000 to as many as 35 million 40grain, .22 caliber long rifle bullets.”²³² Earhart was subsequently executed.²³³

2. Post-Report Developments

Much of the FBI testimony rested on a database, which the Bureau had built up over the course of many years. Although the NAS committee frequently asked for this data during its year-long investigation, the FBI did not turn over the data until it was too late to include an analysis of the information in its report.²³⁴ The two statisticians who served on the NAS committee later wrote that their subsequent inspection of the data “identified several peculiarities.”²³⁵ First, the database was incomplete. The FBI claimed to have a “complete data file” of some 71,000+ measurements but only 64,869 were turned over. Moreover, only measurements made by ICP-AES were included; a different analytical method, NAA, had been used before 1997. Both techniques measured the same elements, and therefore the results from either technique would have been suitable for comparison. Further, the numbering system for the bullets was “highly inconsistent and rather unexpected,” suggesting that some bullet measurements had been deleted.²³⁶ Additionally, “a rough investigation of the measurement

here. Now, bullets that are the same composition can also be found in other boxes of ammunition, but it’s most likely those boxes would have been manufactured at the same place on or about the same date.”). *But see* testimony of Charles Peters, FBI examiner, *Commonwealth v. Wilcox*, Kentucky, Feb. 28, 2002 (*Daubert* hearing: “We have never testified, to my knowledge, that that bullet came from that box. We’d never say that. All we are testifying is that bullet, or that victim fragment or something, the bullet, either came from that box or the *many* boxes that were produced at the same time.” Transcript at 1-2) (emphasis added).

²³² NATIONAL RESEARCH COUNCIL, *supra* note 228, at 6.

²³³ *See* Death Penalty Information Center, Searchable Database of Executions, <http://www.deathpenaltyinfo.org/views-executions> (search for “Earhart” under “Find Person” search box) (last visited Feb. 15, 2017).

²³⁴ *See* Cliff H. Spiegelman & Karen Kafadar, *Data Integrity and the Scientific Method: The Case for Bullet Lead Data as Forensic Evidence*, 19:2 CHANCE 16, 22 (2006) (“During the open sessions of the committee meetings, the FBI claimed to have a ‘complete data file’ of some 71,000+ measurements. Following repeated requests from the Committee, the FBI submitted at its last meeting a CD-ROM that contained two data files with a combined total of 64,869 bullet (not 71,000+) measurement records. . . . This data set could not be analyzed in time for the release of the report . . .”).

²³⁵ *Id.*

²³⁶ *Id.* (“[T]he numbering system of the bullets was highly inconsistent and rather unexpected, *e.g.*, the bullets from a suspect in a particular case might be numbered Q13A, Q13B, Q13C, Q14A, Q14B, Q14C, . . . , leading one to wonder what happened to bullets Q01, Q02, . . . , Q12.”). Other illustrations of incomplete data were noted: “[W]hile most of the bullets indicated 3 measurements, about 30 bullets had six or more measurements.” *Id.* “[O]nly about 50% of the bullets in this data set were identified as having come from one of the four major bullet

error indicated many measurement errors that exceeded the FBI's claimed analytical precision of 2-5%.”²³⁷ Finally, “only 15% of the 1079 cases listed in these two files had measurements from [National Institute of Standards and Technology] . . . making it impossible to determine the frequency of ‘matches’” in some cases.²³⁸ Accordingly, the “missing data and the inconsistent precisions” undermined the Bureau’s public claims.²³⁹ These authors were puzzled by the FBI’s failure to disclose data: “The scientific method is important for science generally; forensic science is no exception. . . . [T]he evidence in this paper suggest that, at least for [CBLA], forensic science failed in the requirement to share the material, methods and data to reach conclusions with the scientific community.”²⁴⁰

The FBI’s response to the NAS report was also disconcerting. The Bureau quickly put out a press release, obscuring the report’s findings.²⁴¹ The release highlighted the committee’s conclusion that the FBI was using appropriate instrumentation and suitable elements for comparison. Yet, these aspects of CBLA were never seriously questioned. Rather, the interpretation of the data was disputed. Only one sentence in the press release addressed this critical issue: “Recommendations by the [NAS] include suggestions to improve the statistical analysis, quality control procedures, as well as expert testimony.”²⁴² The news media read the report quite differently — e.g., “Study Shoots Holes in Bullet Analysis By FBI,”²⁴³ “Report Finds Flaws,”²⁴⁴ “Panel Questions FBI Bullet Analysis,”²⁴⁵ and “Report Questions the Reliability of an F.B.I. Ballistics Test.”²⁴⁶

The Bureau also included the following passage in the press release: “The

manufacturers in the United States (Cascade Cartridge, Inc.; Federal; Remington; Winchester); the ‘complete data file’ of 71,000 bullets may yield a higher proportion of bullets from these four manufacturers.” *Id.*

²³⁷ *Id.*

²³⁸ *Id.*

²³⁹ *Id.*

²⁴⁰ *Id.* at 22-23.

²⁴¹ Department of Justice, FBI News Release, Feb. 10, 2004.

²⁴² *Id.*

²⁴³ Maurice Possley, *Study Shoots Holes in Bullet Analysis By FBI*, CHICAGO TRIB., Feb. 11, 2004, at 14.

²⁴⁴ Charles Pillar, *Report Finds Flaws in FBI Bullet Analysis; Changes are Proposed for the Technique Often Cited in Expert Testimony in Criminal Trials*, L.A. TIMES, Feb. 11, 2004, at 12.

²⁴⁵ Randolph E. Schmid, *Panel Questions FBI Bullet Analysis*, ASSOCIATED PRESS, Feb. 10, 2004.

²⁴⁶ See also Eric Lichtblau, *Report Questions the Reliability of an F.B.I. Ballistics Test*, N.Y. TIMES, Feb. 11, 2004, at 22.

basis of bullet lead compositional analysis is supported by approximately 50 peer-reviewed articles found in scientific publications beginning in the early 1970's. Published research and validation studies have continued to demonstrate the usefulness of the measurements of trace elements within bullet lead."²⁴⁷ In contrast, the NAS report pointed out that there were "very few peer-reviewed articles on homogeneity and the rate of false positive matches" and "outside reviews have only recently been published."²⁴⁸

Over a year later, the FBI discontinued CBLA testing²⁴⁹ and issued another (and similar) press release. Once again, the release minimized the problems, citing the following reason for its decision: "While the FBI Laboratory still firmly supports the scientific foundation of bullet lead analysis, given the costs of maintaining the equipment, the resources necessary to do the examination, and its relative probative value, the FBI Laboratory has decided that it will no longer conduct this exam."²⁵⁰ Nevertheless, a month earlier, Dwight Adams, the laboratory director, had written a private memorandum to the FBI Director specifying different reasons for abandoning the technique, including the following comments: (1) "We cannot afford to be misleading to a jury" and (2) "We plan to discourage prosecutors from using our previous results in future cases."²⁵¹ Neither concern was reflected in the press release.

In the wake of the NAS report, several state courts excluded CBLA evidence.²⁵² Surprisingly, the FBI supplied affidavits in several cases supporting

²⁴⁷ FBI News Release, *supra* note 243.

²⁴⁸ NATIONAL RESEARCH COUNCIL, *supra* note 228, at 100.

²⁴⁹ Eric Lichtblau, *F.B.I. Abandons Disputed Test for Bullets From Crime Scene*, N.Y. TIMES, Sept. 2, 2005, at A12.

²⁵⁰ Department of Justice, FBI News Release, Sept. 1, 2005.

²⁵¹ John Solomon, *FBI's Forensic Test Full of Holes*, WASH. POST, Nov. 18, 2007, at A1.

²⁵² See *Ragland v. Commonwealth*, 191 S.W.3d 569, 580 (Ky. 2006) (noting that "[i]f the FBI Laboratory that produced the CBLA evidence now considers such evidence to be of insufficient reliability to justify continuing to produce it, a finding by the trial court that the evidence is both scientifically reliable and relevant would be clearly erroneous"); *Clemons v. State*, 896 A.2d 1059, 1070, 1078 (Md. 2006) ("CBLA is not admissible under the *Frye-Reed* standard because it is not generally accepted within the scientific community as valid and reliable."; "Based on the criticism of the processes and assumptions underlying CBLA, we determine that the trial court erred in admitting expert testimony based on CBLA because of the lack of general acceptance of the process in the scientific community."); *State v. Behn*, 868 A.2d 329, 331 (N.J. Super. Ct. 2005) (finding the technique was "based on erroneous scientific foundations").

But see Commonwealth v. Fisher, 870 A.2d 864, 871 (Pa. 2005) ("The CBLA evidence, at best, established a possible connection between Appellant and the bullets recovered from the victim's body."). See also *United States v. Davis*, 406 F.3d 505, 509 (8th Cir. 2005) ("Davis's trial counsel cannot be said to be ineffective for failing to challenge the FBI's methodology on a

prosecutors' efforts to sustain convictions based on the technique. In one affidavit, the FBI cited the NAS report but failed to mention that the report had faulted the Bureau's statistical methods. The chair of the NAS committee criticized the affidavit because it did "not discuss the statistical bullet-matching technique, which is key and probably the most significant scientific flaw found by the committee."²⁵³ The affidavit was also misleading because it estimated that the maximum number of .22-caliber bullets in a batch of lead was 1.3 million, when the NAS committee found that the number could be as high as 35 million.²⁵⁴

On November 18, 2007, *60 Minutes* aired a segment on CBLA.²⁵⁵ In an interview, the FBI lab director, now retired, acknowledged that testimony about boxes was "misleading and inappropriate."²⁵⁶ That broadcast, along with a *Washington Post* investigation, questioned the FBI's response to the NAS report. The main problem was that only the FBI had records of all the cases in which its experts had testified, and the Bureau had declined to disclose the names of those cases.²⁵⁷ Instead, the Bureau relied on the NAS report, its own press releases, and pro forma letters sent to prosecution and defense organizations to notify defendants. This method of communication was grossly inadequate because the letters neither highlighted the problem, nor its significance.²⁵⁸ A few days after the *60 Minutes* expose, Senator Patrick Leahy, the Chairman of the Senate Judiciary Committee, sent a letter to the FBI Director noting that the Bureau's letters gave "the false impression that these discredited tests had continuing reliability."²⁵⁹

* * *

basis that was not advanced by the scientific community at the time of trial.").

²⁵³ Solomon, *supra* note 243 (quoting Ken MacFadden).

²⁵⁴ *Id.*

²⁵⁵ *60 Minutes: Evidence of Injustice* (CBS television broadcast, Nov. 18, 2007).

²⁵⁶ *Id.*

²⁵⁷ Solomon, *supra* note 251, at A1 ("Hundreds of defendants sitting in prisons nationwide have been convicted with the help of an FBI forensic tool that was discarded more than two years ago. But the FBI lab has yet to take steps to alert the affected defendants or courts, even as the window for appealing convictions is closing . . .").

²⁵⁸ The Innocence Network and the National Association of Criminal Defense Lawyers formed a task force and worked with the FBI to contact defense attorneys and convicts. See Vesna Jaksic, *Faulty Bullet-Test Cases Finding Way to Court*, NAT'L L.J., Feb. 25, 2008 ("The task force is lining up pro bono commitments from several law firms to handle the cases.").

²⁵⁹ John Solomon, *Leahy Pursues Forensic-Test Answers; Attorney General Is Told to Prepare For Senate Inquiry*, WASH. POST, Nov. 22, 2007, at A2 (quoting). Leahy also wrote: "The new revelations about bullet-lead analysis are just the latest examples of the Department's inadequate efforts to ensure that sound forensic testing is utilized to the maximum extent to find the guilty rather than merely obtain a conviction. Punishing the innocent is wrong and allows the guilty party to remain free." *Id.*

Here, the flaws are many: Lack of foundational research, failure to make a database available to outside scientists, and ignoring the FBI's own protocols by presenting inconsistent and misleading testimony. Moreover, the reluctance to confess error and take timely corrective action violated basic scientific norms. After decades of use, a federal district court in 2003 excluded CBLA evidence under the *Daubert* standard²⁶⁰ for the first time.

F. Arson Investigations

For decades arson investigators came from the “old school” of investigators — those who used intuition and a number of rules of thumb to determine whether a fire was incendiary. Critics of this approach complained that it lacked a scientific foundation. Rather, it was based on folklore that had been passed down from generation to generation — without any empirical testing.²⁶¹ As early as 1977, a government report noted that common arson indicators had “received little or no scientific testing” and “[t]here appears to be no published material in the scientific literature to substantiate their validity.”²⁶² Through the 1980s, proponents of a science-based approach to arson investigations waged an uphill battle, finally winning a major victory in 1992 when the National Fire Protection Association (NFPA) published its *Guide for Fire and Explosion Investigations* (NFPA 921).²⁶³

1. *Willingham Case*

Although NFPA 921 would subsequently become the bible in arson investigations,²⁶⁴ it was published weeks after Cameron Todd Willingham was convicted for the arson-murders of his young children. Willingham, who was executed twelve years later, is the poster-boy for junk science in arson

²⁶⁰ United States v. Mikos, No. 02 CR 137, 2003 WL 22922197 (N.D. Ill. Dec. 9, 2003).

²⁶¹ See JOHN J. LENTINI, SCIENTIFIC PROTOCOLS FOR FIRE INVESTIGATION ch. 8 (2006) (discussing myths of arson investigations).

²⁶² J.F. BOUDREAU ET AL., NATIONAL INSTITUTE OF LAW ENFORCEMENT AND CRIMINAL JUSTICE, LAW ENFORCEMENT ASSISTANCE ADMINISTRATION, U.S. DEP'T OF JUSTICE, ARSON AND ARSON INVESTIGATION: A SURVEY AND ASSESSMENT (1977).

²⁶³ NATIONAL FIRE PROTECTION ASSOCIATION, GUIDE FOR FIRE AND EXPLOSION INVESTIGATION (1992) [NFPA 921]. The NFPA promotes fire prevention and safety. The most recent edition of NFPA 921 was published in 2011.

²⁶⁴ See United States v. Hebshie 754 F. Supp. 2d 89, 111 n.39 (D. Mass. 2010) (NFPA 921 “is widely accepted as the standard guide in the field of fire investigation.”); Thomas M. May, *Fire Pattern Analysis, Junk Science, Old Wives Tales, and Ipse Dixit: Emerging Forensic 3D Imaging Technologies to the Rescue?*, 16 RICHMOND J.L. & TECH. 1, 5 (2010) (noting that NFPA 921 has “become the *de facto* national standard for fire scene examination and analysis”).

investigations.²⁶⁵

At trial Deputy Fire Marshall Vasquez testified that “[t]he fire tells a story. I am just the interpreter. . . . And the fire does not lie. It tells me the truth.”²⁶⁶ He told the jury that he had found twenty arson so-called “indicators” during his post-fire investigation of Willingham’s house.²⁶⁷ One indicator was a low burning fire.²⁶⁸ “All fire goes up,” Vasquez testified.²⁶⁹ Thus, burn patterns on the lower walls and floor suggested that an accelerant was used.²⁷⁰ This common-sense notion, however, has its limitations, especially when a fire occurs in a contained area, such as a house with its windows shut. Due to buoyancy, a thermal plume initially rises once a fire is ignited. As the fire continues, the plume reaches the ceiling, which causes it to spread outward towards the walls. When it reaches the walls, the combustion products press down from the ceiling creating an upper level, which continues to increase in depth and temperature. Eventually thermal radiation replaces convection as the principal method of heat transfer. At this point, every combustible surface in the room will spontaneously burst into flames. This transition phenomenon, known as the onset of “flashover,” can occur within minutes. After flashover, the entire room is burning, including the lower walls and floor. Flashover, according to one authority, is the point at which the fire transitions from a “fire in a room” to a “room on fire.”²⁷¹ At trial, prosecution

²⁶⁵ See *Frontline: Death by Fire* (PBS television broadcast Oct. 19, 2010); David Grann, *Trial by Fire: Did Texas Execute an Innocent Man?*, NEW YORKER, Sep. 7, 2009, at 63; Michael Hall, *False Impressions*, TEX. MONTHLY, Jan. 2016, at 7 (“The 893-page report, released in April 2011, was anticlimactic for people looking for proof that Texas had executed an innocent man.”); Steve Mills & Maurice Possley, *Texas Man Executed on Disproved Forensics: Fire that Killed His 3 Children Could Have Been Accidental*, CHI. TRIB., Dec. 9, 2004, at C1 (“Arson investigators in Texas have relied on old wives’ tales and junk science to send men to prison, and perhaps even the death chamber, top experts on fire behavior say.”).

²⁶⁶ Transcript, *State v. Willingham*, No. 24240-CR (13th Dist., Tex. 1991), vol. XI, at 244 [hereinafter *Willingham transcript*], *aff’d*, *Willingham v. State*, 897 S.W.2d 351, 354 (Tex. Crim. App. 1995).

²⁶⁷ A second expert’s testimony essentially tracked Vasquez’s.

²⁶⁸ Vasquez testified that there was “char burning, like, for example, this is the bottom here. It’s burned down here at the bottom. That is an indicator in my investigation of an origin of fire because it’s the lowest part of the fire.” *Willingham transcript*, *supra* note 266, vol. XI, at 239. See also *Willingham*, 897 S.W.2d at 354 (“An expert witness for the State testified that the floors, front threshold, and front concrete porch were burned, which only occurs when an accelerant has been used to purposely burn these areas. This witness further testified that this igniting of the floors and thresholds is typically employed to impede firemen in their rescue attempts.”).

²⁶⁹ *Willingham transcript*, *supra* note 266, vol. XI, at 232.

²⁷⁰ “So when I found that the floor is hotter than the ceiling, that’s backwards, upside down. It shouldn’t be like that. The only reason that the floor is hotter is because there was an accelerant.” *Id.* at 256.

²⁷¹ LENTINI, *supra* note 261, at 68-70.

witnesses acknowledged that there was an explosion.²⁷² Consequently, a low burning fire is not necessarily indicative of an incendiary origin.

Moreover, some of Vasquez's other "indicators" — splotchy looking areas that he called "puddle configurations" and "pour patterns" — are present after flashover in an accidental fire.²⁷³ Similarly, additional indicators, such as alligatoring (large shiny charred blisters on burned wood), are also explained by flashover. This phenomenon also accounts for another fact that Vasquez thought incriminatory. Willingham told investigators that he had attempted to save his daughters, but the heat was too great and he was forced to run from the house without shoes. Willingham's feet were not burned, and in Vasquez's mind, the burn debris on the floor made that impossible.²⁷⁴ However, if Willingham left his home before flashover, his feet would not have been burned.

Charring under an aluminum threshold of an interior door provided still another clue. Here, again, this may occur in a flashover. Other perceived indicators — melted bed springs,²⁷⁵ multiple points of origins,²⁷⁶ and brown stains

²⁷² See Willingham transcript, *supra* note 266, vol. XI, at 75 ("The windows, the electricity started crackling and popping, and the top of the well — well, I was facing the side of the house, and it just blew out. The flames just blew out. . . . All the windows and the front room was engulfed.") (testimony of Dianne Barbe); *id.* at 96 ("We was running towards the house, me and my mother, we was fixing to go and try to get in, and that's when it was an explosion.") (testimony of Dianne Barbe). Vasquez mentioned flashover in his testimony (*id.* vol. XII, at 47-48), but he does not appear to understand its implications.

²⁷³ According to Vasquez, a burn trailer was etched on the floor. Willingham transcript, *supra* note 266, vol. XI, at 244 ("You can see that on the burnt patterns on this puddle configuration on Exhibit No. 36. This is a strong indicator of a liquid.").

²⁷⁴ "There was fire on the floor. . . . He had no injuries on his feet." *Id.* at 267.

²⁷⁵ "[T]he springs were burned from underneath. This indicates there was a fire under this bed because of the burn underneath the bed." *Id.* at 241.

²⁷⁶ "Multiple areas of origin indicate — especially if there is no connecting path, that they were intentionally set by human hands." Willingham transcript, *supra* note 266, at 255. There are two problems here. First, the fire scene did not exhibit multiple origins, according to independent experts. DOUGLAS CARPENTER ET AL., REPORT ON THE PEER REVIEW OF THE EXPERT TESTIMONY IN THE CASES OF STATE OF TEXAS V. CAMERON TODD WILLINGHAM AND STATE OF TEXAS V. EARNEST RAY WILLIS 11-12 (2006). Second, even if the fire scene had shown multiple points of origin, this would not necessarily indicate an intentional fire. LENTINI, *supra* note 261, at 461-62.

on a concrete floor²⁷⁷ — were also consistent with an accidental blaze.²⁷⁸ Vasquez also relied on the presence of “crazed glass,” which are spider-web patterns on the windows as an indication of arson.²⁷⁹ It was long believed that crazed glass resulted from a fire that burned fast and hot — i.e., one fueled by a liquid accelerant. Yet, subsequent research demonstrated that crazing occurs from rapid cooling when water from fire hoses is sprayed on heated windows.²⁸⁰

In retrospect, the most damning piece of evidence involved one of the numerous debris samples submitted for laboratory analysis.²⁸¹ It came from an area near the front door and was the *only* sample that tested positive for a chemical commonly used in charcoal lighter fluids. Nevertheless, this finding can be explained by the fact that a charcoal grill and lighter fluid were on the front porch at the time of the fire.²⁸² In fact, the negative results from the other samples supported Willingham’s case.²⁸³

Numerous nationally-recognized experts reviewed the arson testimony presented at Willingham’s trial and found it seriously flawed. The first examination of the record by an independent expert was submitted to the governor and the Board of Pardons and Parole days before Willingham’s execution. It

²⁷⁷ Willingham transcript, *supra* note 266, vol. XI, at 248-49. Fire experts reviewing the evidence from Willingham’s trial pointed out that “[t]he behavior of concrete in fires, including the development of various colors, has been extensively studied.” CARPENTER ET AL., *supra* note 276, at 18. These experts concluded that there is simply “no scientific basis for Mr. Vasquez’s statement about the brown discoloration being an indication of the presence of accelerants.” *Id.*

²⁷⁸ Vasquez’s testimony also demonstrated other misconceptions. A common one is that arson fires burn hotter and faster than “normal” fires: “You know, it makes the fire hotter. It’s not a normal fire.” Willingham transcript, *supra* note 266, vol. XI, at 249. However, the temperature of burning wood and burning gasoline are nearly identical, so to claim that a fire using liquid accelerants burns “hotter” than a wood fire is wrong. LENTINI, *supra* note 254, at 465.

²⁷⁹ “The pieces of broken window glass on the ledge of the north windows to the northeast bedroom disclosed a crazed ‘spider webbing’ condition. This condition is an indication that the fire burned fast and hot.” CARPENTER ET AL., *supra* note 278, at 18 (citing Vasquez’s written report on the Willingham fire at 4).

²⁸⁰ LENTINI, *supra* note 261, at 439 (“It is unclear why anyone ever thought that crazing of glass indicated rapid heating.”).

²⁸¹ In closing argument, the defense counsel referred to a “dozen samples.” Willingham transcript (vol. XIII), *supra* note 266, at 20.

²⁸² *Id.* at 15 (although photographs show a grill, Vasquez apparently did not know of the grill’s presence); *id.* at 16 (acknowledging that a fire-damaged charcoal lighter fluid container was found on the front porch).

²⁸³ The prosecutor would later say that he “‘never did understand why they weren’t able to recover’ positive tests in these parts.” Grann, *supra* note 265, at 61. At trial, he argued that the “liquid burned away in that destructive madness created by Cameron Todd Willingham.” Willingham transcript, *supra* note 266, vol. XIII, at 45.

concluded: “On first reading, a contemporary fire origin and cause analyst might well wonder how anyone could make so many critical errors in interpreting the evidence.”²⁸⁴ Nevertheless, a stay was denied, and Willingham was put to death. Subsequent evaluations agreed that the trial evidence was junk science. For example, five independent experts prepared a forty-three page report, finding that “each and every one of the indicators relied upon have since been scientifically proven to be invalid.”²⁸⁵

In May 2006, the Innocence Project petitioned the Texas Forensic Science Commission (TFSC) to review the arson testimony in Willingham’s and Ernest Ray Willis’ cases.²⁸⁶ The TFSC is not authorized to determine guilt or innocence. Instead, the Innocence Project argued that the State Fire Marshall Office should have reinvestigated arson cases in which its experts testified after NFPA 921 was published in 1992 — a full twelve years before Willingham’s execution.²⁸⁷ TFSC retained its own independent consultant, Dr. Craig Beyler, another nationally-recognized expert, to review the arson evidence. His fifty-one page report dissected the expert testimony, concluding:

The investigations of the Willis and Willingham fires did not comport with either the modern standard of care expressed by NFPA 921, *or the standard of care expressed by fire investigation texts and papers in the period 1980-1992*. The investigators had poor understandings of fire science and failed to acknowledge or apply the contemporaneous understanding of the limitations of fire indicators. Their methodologies did not comport with the scientific method or the process of elimination. A finding of arson could not be sustained based upon the standard of care expressed by NFPA, or the standard of care expressed by fire investigation texts and papers in the period 1980-1992.²⁸⁸

Once Beyler’s report became public, a political firestorm erupted, and the

²⁸⁴ Report of Dr. Gerald Hurst, In re Cameron Todd Willingham, Trial Court No. 24, 4670(B), District Court, 366th Jud. Dist., Navarro County, Tex., Feb. 13, 2004.

²⁸⁵ CARPENTER ET AL., *supra* note 282.

²⁸⁶ The expert evidence in both cases was comparable, but Willis was lucky. His death penalty conviction was overturned on procedural grounds, and the prosecutor subsequently refused to reindict him after Dr. Hurst wrote the same type of critical report in Willis’s case that he had written in Willingham’s. Willis, who had spent seventeen years on death row, was subsequently exonerated on actual innocence grounds. See Mary Alice Robbins, *New-York Based Innocence Project Attacks Texas Arson Convictions*, 22 TEX. LAWYER, May 8, 2006.

²⁸⁷ See Letter from Innocence Project to Texas Forensic Science Comm’n (Aug. 20, 2010).

²⁸⁸ CRAIG L. BEYLER, ANALYSIS OF THE FIRE INVESTIGATION METHODS AND PROCEDURES USED IN THE CRIMINAL ARSON CASES AGAINST EARNEST RAY WILLIS AND CAMERON TODD WILLINGHAM, Aug. 17, 2009, at 51 (emphasis added).

governor, who was in the midst of a reelection battle, abruptly replaced commission members three days before a meeting scheduled to consider the Beyler report.²⁸⁹ The newly-appointed chair, a prosecutor, promptly cancelled the meeting,²⁹⁰ raising the specter of a cover-up.²⁹¹ Next, the Attorney General issued an opinion ruling that the TFSC lacked jurisdiction over cases decided before its creation.²⁹² The State Fire Marshall vigorously defended its investigation.

The TFSC eventually produced a report — one that did not directly deal with the Willingham and Willis cases. Nevertheless, the report’s recommendations and statements indicated that the Willingham arson investigation was seriously flawed. Its first recommendation was “that fire investigators adhere to the standards of NFPA 921.”²⁹³ In addition, the report reviewed a number of arson indicators that were used in the Willingham and Willis cases. Citing Vasquez’s testimony, the report undermined his opinions concerning (1) V-patterns as an indicator of origin, (2) pour patterns, (3) low/deep burning, (4) multiple separate points of origin, (5) spalling, (6) burn intensity, and (7) crazed glass.²⁹⁴ It also observed that “testimony, such as Vasquez’s response to a question regarding Willingham’s state of mind, is an example of the type of testimony that experts should avoid as falling outside of their field of

²⁸⁹ See Christy Hoppe, *Perry Defends Removing 3: He Says He’s Following Protocol, but Critics Believe He’s Derailing Arson Inquiry*, DALLAS MORNING NEWS, Oct. 2, 2009, at 3A; Mary Alice Robbins, *Fired Up: Changes Sought for Texas Forensic Science Commission at Center of Heated Controversy*, 25 TEX. LAWYER, Nov. 9, 2009 (“[Former Commissioner] Levy says he believes ‘things went south’ for the commission after [former Chair] Bassett released Beyler’s report to the public in August ‘as he was required by law to do.’”). The meeting was scheduled for October 2, 2009.

²⁹⁰ Hoppe, *supra* note 289 (noting that the new chair was “known as one of the toughest law-and order prosecutors in the state”).

²⁹¹ See Jennifer Emily, *Texas Forensic Science Commission Refuses to End Inquiry into Willingham Arson Case*, DALLAS MORNING NEWS, Sept. 18, 2010 (“Perry’s replacements were seen by some as a political maneuver intended to change the outcome of the commission’s decision.”); Christy Hoppe, *Perry Ousts Officials Before Arson Hearing: He’s Assailed as New Chair Delays Session on Flawed Case that Led to Execution*, DALLAS MORNING NEWS, Oct. 1, 2009, at 1A; David Mann, *Fire and Innocence*, TEX. OBSERVER, Dec. 3, 2009 (“Then in late September, Perry booted three members off of the Texas Forensic Science Commission, which was investigating the Willingham and Willis cases, just three days before a crucial hearing on scientists’ findings. Perry’s new appointees promptly canceled the hearing and have yet to reschedule it. Even conservative commentators cried cover-up, suggesting that Perry, in a tough battle for re-election, was trying to subvert an investigation that might prove he oversaw the execution of an innocent man.”).

²⁹² Letter from Greg Abbott, Texas Attorney General, to Texas Forensic Science Comm’n (Aug. 2011).

²⁹³ REPORT OF THE TEXAS FORENSIC SCIENCE COMM’N, WILLINGHAM/WILLIS INVESTIGATION 39 (April 15, 2011).

²⁹⁴ *Id.* at 21-28.

expertise.”²⁹⁵ The report even encouraged lawyers to “aggressively pursue admissibility hearings in arson cases.”²⁹⁶

3. *Han Tak Lee Case*

Unfortunately, Willingham’s case was not an outlier. In the 1989 trial of Han Tak Lee,²⁹⁷ the expert also relied on the old “myths” to declare the fire incendiary: (1) greater intensity and heat, (2) burn patterns, (3) alligatoring, (4) melted metal in bed frames, and (5) crazed glass.²⁹⁸ In addition, the investigation was “hobbled by an incomplete and inaccurate understanding” of flashover. After serving twenty-five years, Lee was released from prison in 2015.²⁹⁹

3. *National Fire Protection Association Guidelines*

After the publication of NFPA 921 in 1991, the kind of testimony presented in the Willingham and Lee cases should have vanished from the courtroom. But arson investigators balked. According to one expert, “[t]he initial response to NFPA 921 in the fire investigation community was overwhelmingly negative.”³⁰⁰ *Babick v. Berghuis*³⁰¹ is illustrative. In that case, Andrew Babick was convicted of arson-murder for a 1995 house fire and was sentenced to two terms of life imprisonment without the possibility of parole. He later sought habeas relief, claiming ineffective assistance of counsel and prosecutorial misconduct. In 2010, the Sixth Circuit rejected these claims.

However, in dissent, Judge Merritt chastised the defense attorney for not

²⁹⁵ *Id.* at 36.

²⁹⁶ *Id.* at 48.

²⁹⁷ Lee petitioned for a writ of habeas corpus in 2010, based in part on “inaccurate and unreliable evidence.” *Han Tak Lee v. Tennis*, No. 4:CV-08-1972, 2010 WL 3812160, at *2 (M.D. Pa. Sept. 22, 2010). Although the district court denied Lee’s petition, the Third Circuit reversed. *Han Tak Lee v. Glunt*, 667 F.3d 397, 407-08 (3d Cir. 2012) (“If Lee’s expert’s independent analysis of the fire scene evidence — applying principles from new developments in fire science — shows that the fire expert testimony at Lee’s trial was fundamentally unreliable, then Lee will be entitled to federal habeas relief on his due process claim.”).

²⁹⁸ *Han Tak Lee v. Tennis*, Civil No. 4:08–CV–1972, 2014 WL 3894306 (M.D. Pa. June 13, 2014) (magistrate report), *aff’d sub. nom.*, *Han Tak Lee v. Houtzdale*, 798 F.3d 159 (3d Cir. 2015).

²⁹⁹ Mark Hansen, *Badly Burned: Long-Held Beliefs About Arson Science Have Been Debunked After Decades of Misuse and Scores of Wrongful Convictions*, 101 A.B.A. J. 37 (Dec. 2015).

³⁰⁰ *Id.*

³⁰¹ 620 F.3d 571 (6th Cir. 2010). See generally Marc Price Wolf, *Habeas Relief from Bad Science: Does Federal Habeas Corpus Provide Relief for Prisoners Possibly Convicted on Misunderstood Fire Science?*, 10 MINN. J.L. SCI. TECH. 213 (2009).

contesting the arson evidence in “this strange junk science case.”³⁰² One prosecution expert testified that: (1) char marks on the porch were evidence of an accelerant, (2) a “line of demarcation” in a burn pattern on a carpet was “suspicious” because “it should not have burned the carpeting on these jagged edges,” and (3) the burns were “not normal” and were “unnatural.”³⁰³ Another prosecution expert stated that “low burning” and other “unnatural” patterns indicated the presence of an accelerant. Both experts “testified — in direct contrast to the NFPA guide — that they were so confident in their reading of burn patterns that the absence of any laboratory confirmation of accelerant had no effect on their testimony.”³⁰⁴

4. Dog-sniff Evidence

More alarming, in Judge Merritt’s view, was dog-sniff evidence. The NFPA guide provides: “Research has shown that canines have been alerted to pyrolysis products that are not produced by an ignitable liquid” and a positive canine alert without laboratory confirmation “should not be considered validated.”³⁰⁵ The lab tests had not detected accelerants in the house debris. Yet, a dog handler testified that “his dog, Samantha, was ‘1000 times’ more effective at detecting fire starters or liquid accelerants than a laboratory test on burnt material.”³⁰⁶ In short, the “jury was misled into trusting Samantha over the arson forensic lab.”³⁰⁷

A more recent arson-dog case involved James Hebshie, who was convicted of arson and mail fraud in 2006. A federal district court granted his habeas petition based on ineffective assistance of counsel grounds.³⁰⁸ In the court’s view, had a *Daubert* hearing been requested on the canine evidence, there was a “‘reasonable probability’ that the Court would have excluded the canine testimony or severely limited it.”³⁰⁹ Without a challenge from the defense, the dog

³⁰² 620 F.3d at 580.

³⁰³ *Id.* at 581 (quoting transcript).

³⁰⁴ *Id.*

³⁰⁵ NFPA 921, *supra* note 263, § 16.5.4.7 (describing the role of canine investigations as “assisting with the location and collection of samples” for laboratory testing).

³⁰⁶ *Babick*, 620 F.3d at 580.

³⁰⁷ *Id.* See also *United States v. Myers*, No. 3:10–00039, 2010 WL 2723196 (S.D.W.Va. July 8, 2010) (granting motion in limine to prohibit expert testimony of a canine handler because the alert had not been confirmed by lab testing, conflicted with the Fire Guide, and did not meet the *Daubert* standards).

³⁰⁸ *United States v. Hebshie*, 754 F. Supp. 2d 89 (D. Mass. 2010).

³⁰⁹ *Id.* at 124.

handler testified that his dog (Billy) was 97% accurate.³¹⁰ Indeed, the handler testified to “an almost mystical account of Billy’s powers and her unique olfactory capabilities.”³¹¹ The court explained: “[The handler] went on and on about what he understood about Billy, as if his relationship with Billy somehow enhanced the reliability and probative value of the results — that she was unique, that he could ‘read her face,’ that he was with her 365 days a year, that he knew her personality, ‘the way her eyes shifted,’ the ways her ear shifted, etc.”³¹²

The handler focused on one area as the origin of the fire and testified that the dog had not alerted anywhere else on the premises. However, the handler had limited the dog’s access to that one area. In addition, a dog’s failure to alert has no evidential value: “[T]he scientific literature cast doubt on the significance of the dog’s *failure* to alert (false negatives) and even raised concerns about canine ‘proficiency’ testing, concerns counsel never raised.”³¹³ Indeed, the term “accelerant-detection” dog was misleading because the dog is trained to alert to many common materials that are not accelerants; the site of the fire was a convenience store which sold lighter fluid and lighters.

5. *Post-Daubert Cases*

The courts’ response to bogus arson evidence is mixed.³¹⁴ It is not hard to find cases citing discredited arson indicators after *Daubert*, such as pour patterns or puddle configurations,³¹⁵ melted bedsprings,³¹⁶ concrete spalling,³¹⁷ fire load,³¹⁸

³¹⁰ See Michael E. Kurtz et al., *Effect of Background Interference on Accelerant Detection by Canines*, 41 J. FORENSIC SCI. 868 (1996) (discussing the varying levels of reliability in accelerant detection depending on the substance in question and the canine handler); see also *Farm Bureau Mut. Ins. Co. v. Foote*, 14 S.W.3d 512, 518 (Ark. 2000) (affirming the trial court’s exclusion of a canine handler who sought to testify about “the alleged superior ability of his canine partner, Benjamin, to detect the presence of accelerants after a fire . . . [that he could] discriminate between different types of chemicals,” and that he had an accuracy rate of “100%”).

³¹¹ *Hebshie*, 754 F. Supp. 2d at 93.

³¹² *Id.* at 119.

³¹³ *Id.* at 94.

³¹⁴ 1 GIANNELLI ET AL., *supra* note 15, at 1102-03 (“Many appellate cuts continue to routinely accept investigators’ testimony about experientially based generalizations.”) (citing cases).

³¹⁵ See, e.g., *State v. Allen*, No. 22835, 2009 WL 2096295 ¶ 114 (Ohio Ct. App. 2009) (investigator testifying to “an irregular burn pattern on the floor which through all my experience and training it appears to be an irregular pour patterns [sic], an ignitable liquid pour pattern”); *State v. Wolf*, 891 N.E.2d 358, 360 (Ohio Ct. App. 2008) (A firefighter testified “that he observed ‘pour patterns’ located on the floor throughout the mobile home; that the pour patterns are burnt marks that look like puddles that result from ignitable liquids . . . being poured out of containers”); *Colburn v. State*, 990 So. 2d 206, 209-10 (Miss. Ct. App. 2008) (“This pour pattern, [the fire investigator] explained, was indicative of flammable liquid being poured in the area On cross-examination [the fire investigator] did admit that the State Crime Laboratory was unable to identify ignitable liquids in the three debris samples taken from the pour pattern

and “fast and hot” burn.³¹⁹ Decided in 1998, *Michigan Millers Mut. Ins. Corp. v. Benfield*³²⁰ is considered the “first serious challenge to the ‘old school’ of fire investigators.”³²¹ In that case, the Eleventh Circuit ruled that arson testimony “is subject to *Daubert*’s inquiry regarding the reliability of such testimony”³²² and cited NFPA 921.³²³ Yet, a 2011 article on the subject began with the passage: “Fire researchers have shattered dozens of arson myths in recent years. So why do American courts still lag behind?”³²⁴ And a 2013 survey of 586 public sector fire investigators found that some myths endure: “Nearly 40 percent did not know that crazed glass is caused by rapid cooling, not rapid heating. Twenty-three percent think puddle-shaped burns indicate the use of an accelerant. Eight percent still believe that alligator blistering implies that a fire burned fast and hot.”³²⁵

* * *

area.”); *State v. Henderson*, 125 P.3d 1132, 1137 (Mont. 2005) (finding that the trial court “did not err in allowing [a firefighter] to identify in the photographs and diagrams the pour patterns he had observed at the scene”).

³¹⁶ *Simon v. State*, 633 So. 2d 407, 409 (Miss. 1993), *vacated*, *Simon v. Mississippi*, 513 U.S. 956 (1994).

³¹⁷ *See, e.g.*, *State v. Amodio*, 915 A.2d 569, 576 (N.J. Super. Ct. App. Div. 2007) (“They washed the floor and observed areas of spalling in the concrete underneath the door. This was an indication that a flammable liquid had been employed in that area.”); *McCord v. Gulf Guar. Life Ins. Co.*, 698 So. 2d 89, 95 (Miss. 1997) (“The arson investigator ... testified that he found five different areas of spalling and concluded arson to be the cause of the fire.”).

³¹⁸ *See, e.g.*, *Wise v. State*, 719 N.E.2d 1192, 1200 (Ind. 1999) (A fire investigator testified that a fire was intentionally set based on several factors, including that “the fire burned too fast for its fuel load.”); *Carter v. State*, 516 S.E.2d 556, 560 (Ga. Ct. App. 1999) (The fire investigator “deduced there must have been an accelerant or some kind of extra fuel load.”).

³¹⁹ *See, e.g.*, *People v. Klait*, No. 06-000399-FH, 2010 WL 2076956, at *5 (Mich. Ct. App. May 25, 2010) (“[T]hey both testified that they believed, based on the fast and hot nature of the fire, that it was set intentionally.”); *State v. Walters*, 813 P.2d 857, 858 (Idaho 1990) (A fire investigator testified that “it was a hot, fast fire as opposed to a small or as opposed to a slow, smoldering fire, yes, the evidence suggests to me that it was deliberately set.”); *State v. Cutlip*, No. 99-L-149, 2001 WL 687493, at *2 (Ohio Ct. App. 2001) (A fire department lieutenant testified to a list of factors including that “the fire was fast and hot” and “that such observations are typical of a fire started by someone pouring an accelerant and lighting it.”).

³²⁰ 140 F.3d 915, 920 (11th Cir. 1998).

³²¹ John J. Lentini, *The Evolution of Fire Investigations and Its Impact on Arson Cases*, 27 CRIM. JUST. 12, 14 (Spring 2012).

³²² *Benfield*, 140 F.3d at 920.

³²³ *See also Fireman’s Fund Ins. Co. v. Canon U.S.A., Inc.*, 394 F.3d 1054, 1058 (8th Cir. 2005) (holding district court’s exclusion of expert arson evidence proper where experts failed to compare hypothesis to evidence from scene in violation of NFPA 921); *Ind. Ins. Co. v. Gen. Elec. Co.*, 326 F. Supp. 2d 844, 850–51 (N.D. Ohio 2004) (holding that cause-and-origin expert’s failure to properly collect evidence violated NFPA 921).

³²⁴ Douglas Starr, *Up in Smoke*, DISCOVER 36, 37 (Nov. 2011).

³²⁵ Hansen, *supra* note 299, at 42-43.

The Texas Forensic Science Commission’s report did more than the courts to curb flawed arson testimony. Moreover, the number of ineffective assistance of counsel cases, such as *Babick* and *Hebshie*, exposed another glaring defect in the criminal justice system’s capability to evaluate expert testimony. For example, in *Richey v. Bradshaw*,³²⁶ the Sixth Circuit wrote: “The scientific evidence of arson was thus fundamental to the State’s case. Yet Richey’s counsel did next to nothing to determine if the State’s arson conclusion was impervious to attack.”³²⁷ Similarly, in *Dugas v. Coplan*,³²⁸ the First Circuit criticized counsel because his “investigation consisted of his own visual assessment of the fire scene, his conversations with the state’s experts, some limited reading, and his conversations with other defense attorneys after work.”³²⁹ Then, without consulting an arson expert, he mounted a “not arson” defense.

III. FORENSIC SCIENCE RESEARCH

By now it is almost a truism that too many forensic disciplines are not grounded in science — and yet their adherents continue to claim the mantle of science. The NAS report emphasized the “notable dearth of peer-reviewed, published studies establishing the scientific bases and validity of many forensic methods.”³³⁰ Indeed, the co-chair of the NAS committee, Judge Harry Edwards, later stated: “I think that the most important part of our Committee’s Report is its call for real science to support the forensic disciplines.”³³¹ Not surprisingly, the report triggered extensive commentary.³³² One cataloged the numerous ways in which forensic science has failed to develop a research culture³³³ and argued that the “core values” of a scientific culture “are empiricism, transparency, and an ongoing critical perspective.”³³⁴ Another article documented the serious problems that have arisen when the law enforcement controls forensic research.³³⁵

³²⁶ 498 F.3d 344 (6th Cir. 2007).

³²⁷ *Id.* at 362.

³²⁸ 428 F.3d 317 (1st Cir. 2005).

³²⁹ *Id.* at 323.

³³⁰ NAS FORENSIC REPORT, *supra* note 21, at 8.

³³¹ Honorable Harry T. Edwards, The National Academy of Sciences Report on Forensic Sciences: What it Means for the Bench and Bar, Address at Conference of Superior Court of the District of Columbia, Washington, D.C., May 6, 2010, at 7.

³³² See Paul C. Giannelli, *The 2009 NAS Report: A Literature Review*, 48 CRIM. L. BULL. 378 (2012) (listing numerous articles and conferences).

³³³ Jennifer L. Mnookin et al., *The Need for a Research Culture in the Forensic Sciences*, 58 UCLA L. REV. 725 (2011).

³³⁴ *Id.* at 742.

³³⁵ See Paul C. Giannelli, *Daubert and Forensic Science: The Pitfalls of Law Enforcement Control of Scientific Research*, 2011 U. ILL. L. REV. 53.

The NAS report recommended the creation of an independent federal entity (the National Institute of Forensic Sciences) to oversee the field, including a research agenda.³³⁶ If adopted, this proposal would have wrest control of forensic science from law enforcement. According to the report, some federal entities were “too wedded” to the status quo and “have failed to pursue a rigorous research agenda to confirm the evidentiary reliability of methodologies used in a number of forensic science disciplines.”³³⁷ As a result, these “agencies are not good candidates to oversee the overhaul of the forensic science community.”³³⁸ Unfortunately, Congress did not act on this recommendation.

A. National Commission on Forensic Science (2013-17)

To its credit, the DOJ, in partnership with the National Institute of Standards and Technology (NIST), established the National Commission on Forensic Science in 2013. The commission’s task was to enhance the practice and improve the reliability of forensic science.³³⁹ Early on, the commission created a subcommittee on scientific inquiry and research, which undertook the task of reviewing bibliographies of foundational literature that had been compiled by various forensic disciplines.³⁴⁰ The subcommittee quickly concluded that even a “cursory review” of the bibliographies raised serious concerns. One basic problem involved the definition of foundational literature. According to the subcommittee, “[i]n some cases, it was unclear which literature citations are

³³⁶ NAS FORENSIC REPORT, *supra* note 21, at 19 (Recommendation 1(c): “promoting scholarly, competitive peer-reviewed research and technical development in the forensic science disciplines”).

³³⁷ *Id.* at 18.

³³⁸ *Id.* There is little question that the committee was referring to National Institute of Justice and the FBI Laboratory. The report noted that, although both had provided “modest leadership” in forensic science, “neither entity has recognized, let alone articulated, a need for change or a vision for achieving it.” *Id.* at 16. The report also stated: “Neither has the full confidence of the larger forensic science community. And because both are part of a prosecutorial department of the government, they could be subject to subtle contextual biases that should not be allowed to undercut the power of forensic science.” *Id.* Consequently, “advancing *science* in the forensic science enterprise is not likely to be achieved within the confines of DOJ.” *Id.* at 18.

³³⁹ National Commission on Forensic Science, U.S. Department of Justice, <https://www.justice.gov/ncfs> (last visited ___).

³⁴⁰ As a result of the NAS report, an Interagency Working Group — the Research Development Technology and Evaluation (RDT&E) of the National Science and Technology Council’s Subcommittee on Forensic Science was tasked with identifying foundational research forensic sciences. National Science and Technology Council Committee on Science Subcommittee on Forensic Science, May 2, 2014, Office of Science & Technology Policy. The RDT&E committee requested Scientific Working Groups (SWG) with addressing a series of discipline-specific questions. In response, literature compendiums were submitted to the RDT&E committee by several forensic working groups.

crucial to support the foundation of a particular forensic discipline.”³⁴¹ This finding led the subcommittee to define the term: foundational, scientific literature should consist of “original research, substantive reviews of the original research, clinical trial reports, or reports of consensus development conferences.”³⁴² Tellingly, the subcommittee felt compelled to add: “While other forms of dissemination of research and practice (e.g., oral and poster presentations at meetings, workshops, personal communications, editorials, dissertations, theses, and letters to editors) play an important role in science, the open, peer-reviewed literature is what endures and forms a foundation for further advancements.”³⁴³

The subcommittee’s second concern was that “some of the cited literature had not undergone a rigorous peer-review process.”³⁴⁴ Peer review by other members of a forensic discipline is not sufficient.³⁴⁵ Many of the reviewers are not scientists, and there is the problem with role bias. According to the subcommittee, foundational research should be subjected to “rigorous peer review with independent external reviewers to validate the accuracy . . . [and] overall consistency with scientific norms of practice”³⁴⁶ and “published in a journal that is searchable using free, publicly available search engines.”³⁴⁷ With few exceptions, the disciplines considered above have not satisfied these requirements.³⁴⁸

³⁴¹ Nat’l Comm’n on Forensic Sci., U.S. Department of Justice, Views Document on Scientific Literature in Support of Forensic Science and Practice (adopted at NCFS Meeting #5 – January 30, 2015) [hereinafter Views Document on Scientific Literature].

³⁴² *Id.*

³⁴³ *Id.*

³⁴⁴ *Id.*

³⁴⁵ *See* United States v. Crisp, 324 F.3d 261, 274 (4th Cir. 2003) (Michael, J., dissenting) (“Fingerprint examiners, . . . have their own professional publications. . . . But unlike typical scientific journals, the fingerprint publications do not run articles that include or prompt critique or reanalysis by other scientists. Indeed, few of the articles address the principles of fingerprint analysis and identification at all . . .”). *See also* Zabell, *supra* note 155, at 164 (“Although there is a substantial literature on the uniqueness of fingerprints, it is surprising how little true scientific support for the proposition exists.”).

³⁴⁶ Views Document on Scientific Literature, *supra* note 341 (“Published in a journal that maintains a clear and publicly available statement of purpose that encourages ethical conduct such as disclosure of potential conflicts of interest integral to the peer review process.”).

³⁴⁷ *Id.* Other publication requirements include: (1) “Published in a journal or book that has an International Standard Number (ISSN for journals; ISBN for books) and recognized expert(s) as authors (for books) or on its Editorial Board (for journals).” (2) “Published in a journal that is indexed in databases that are available through academic libraries and other services (e.g. JSTOR, Web of Science, Academic Search Complete, and SciFinder Scholar).”

³⁴⁸ Another commission document provided guidance for evaluating scientific literature. Nat’l Comm. on Forensic Sci., Department of Justice, Views Document on Identifying and Evaluating Literature that Supports the Basic Principles of a Forensic Science Method or Forensic Science Discipline (adopted at NCFS Meeting #9 – March 22, 2016). Including:

- Is the problem or hypothesis clearly stated?

Another recommendation, one on technical merit, provides: “All forensic science methodologies should be evaluated by an independent scientific body to characterize their capabilities and limitations in order to accurately and reliably answer a specific and clearly defined forensic question.”³⁴⁹ Significantly, the commission recommended that the NIST be the independent scientific evaluator within the justice system.

B. White House PCAST Report (2016)

Unlike the commission, which had a broad mandate, the White House PCAST report focused only on the validation issue. It took pains to explain the concept of validation, noting that forensic methods must be based on empirical studies and be “repeatable, reproducible, and accurate, at levels that have been measured and are appropriate to the intended application.”³⁵⁰ The report recognized that forensic methods may be either objective or subjective. Foundational validity for objective methods “can be established by studying [and] measuring the accuracy, reproducibility, and consistency of each of its individual steps.”³⁵¹ By definition, this approach is not possible with subjective techniques

-
- Is the scope of the article clearly stated as appropriate (article, case study, review, technical note, etc.)?
 - Is the literature review current, thorough, and relevant to the problem being studied?
 - Does this work fill a clear gap in the literature or is it confirmatory and/or incremental?
 - Are the experimental procedures clear and complete such that the work could be easily reproduced?
 - Are the experimental methods appropriate to the problem?
 - Are the methods fully validated to the necessary level of rigor (fit for purpose)?
 - Are the data analysis and statistical methodology appropriate for the problem, and explained clearly so it can be reproduced?
 - Are the experimental results clearly and completely presented and discussed?
 - Are omissions and limitations to the study discussed and explained?
 - Are the results and conclusions reasonable and defensible based on the work and the supporting literature?
 - Are the citations and references complete and accurate?
 - Are the references original (primary) and not secondary?
 - Are funding sources and other potential sources of conflict of interest clearly stated?

³⁴⁹ Nat’l Comm. on Forensic Sci., Department of Justice, Views Document on Technical Merit Evaluation of Forensic Science Methods and Practices (Adopted at NCFS Meeting #10 - June 21, 2016). Recommendation: Recommendation on Technical Merit Evaluation of Forensic Science Methods and Practice (adopted at NCFS Meeting #11 - September 12, 2016).

³⁵⁰ WHITE HOUSE PCAST REPORT, *supra* note 33, at 4-5. Here, “repeatable” means an examiner reaches the same result when analyzing the same sample. “Reproducible” means that different examiners reach the same result when analyzing the same sample. The term “accurate” means that “an examiner obtains correct results both (1) for samples from the same source (true positives) and (2) for samples from different sources (true negatives).” Finally, “reliability” means repeatability, reproducibility, and accuracy. *Id.* at 47.

³⁵¹ *Id.* at 5.

because they involve significant human judgment. Consequently, validity and reliability for these methods must be based on “black-box studies” (as if a “black box” is in the examiner’s head), in which numerous examiners make decisions on many independent tests in order to determine error rates.³⁵²

Importantly, the report also specified what does *not* qualify as validation: “experience, judgment, good professional practices (such as certification programs and accreditation programs, standardized protocols, proficiency testing, and codes of ethics) cannot substitute for actual evidence of foundational validity and reliability.”³⁵³ Moreover, expressions of *confidence* by individual examiners or a consensus among practitioners about the accuracy cannot substitute for “error rates estimated from relevant studies.” In sum, empirical evidence is the “sine qua non” for establishing foundational validity.³⁵⁴

PCAST also recommended that NIST conduct scientific evaluations of the validity of current and new forensic technologies: “To ensure the scientific judgments are unbiased and independent, such evaluations should be conducted by an agency which has no stake in the outcome.”³⁵⁵

In response, DOJ released a statement criticizing the report — on the day of its release. According to DOJ, the PCAST report “does not mention numerous published research studies which seem to meet PCAST’s criteria for appropriately designed studies providing support for foundational validity. That omission discredits the PCAST report as a thorough evaluation of scientific validity.”³⁵⁶ PCAST, in turn, invited all stakeholders to identify validity studies that it might have overlooked. “DOJ ultimately concluded that it had no additional studies for PCAST to consider.”³⁵⁷ Nor did the more than 400 papers submitted by twenty-six respondents cause PCAST to change its positions. The bottom line remained: “In science, empirical testing is the only way to establish the validity and degree

³⁵² *Id.* at 5-6.

³⁵³ *Id.*

³⁵⁴ *Id.*

³⁵⁵ *Id.* The NAS report considered NIST before recommending an independent agency but rejected the idea because, at that time, NIST had limited ties to forensic science. NAS FORENSIC REPORT, *supra* note 21, at 17.

³⁵⁶ Department of Justice, Comment Letter on PCAST’s Report to the President on Forensic Science in Federal Criminal Courts: Ensuring Scientific Validity of Pattern Comparison Methods (Sept. 20, 2016), <http://www.crime-scene-investigator.net/PDF/fbi-response-to-forensic-science-in-federal-criminal-courts-ensuring-scientific-validity-of-pattern-comparison-methods.pdf>.

³⁵⁷ PRESIDENT’S COUNCIL OF ADVISORS ON SCIENCE AND TECHNOLOGY, AN ADDENDUM TO THE PCAST REPORT ON FORENSIC SCIENCE IN CRIMINAL COURTS 3 (Jan. 6, 2017), https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensics_addendum_finalv2.pdf.

of reliability of such an empirical method. Fortunately, empirical testing of empirical methods is feasible. There is no justification for accepting that a method is valid and reliable in the absence of appropriate empirical evidence.”³⁵⁸ However, most prior studies use “closed-set design.” In these studies, “the correct source of each questioned sample is always present; studies using the closed-set design have underestimated the false-positive and inconclusive rates by more than 100-fold.”³⁵⁹

IV. CONCLUSION

This article examined the courts’ systemic failure in criminal cases to fulfill its “gatekeeper”³⁶⁰ function under *Daubert*. The courts, of course, function as part of a justice system that relies on attorneys to discredit erroneous or overstated testimony. According to the *Daubert* Court, “[v]igorous cross-examination, presentation of contrary evidence, and careful instruction on the burden of proof are the traditional and appropriate means of attacking shaky but admissible evidence.”³⁶¹ After the release of the NAS report, some commentary focused on defense counsel’s incompetence: “Criminal defense lawyers . . . are supposed to be the people who recognize bogus expert claims, challenge them, move to get them excluded, and undermine those that survive exclusion by knowledgeable, thorough, and telling cross-examination. On the whole, they don’t do any of these things very well.”³⁶² Judge Nancy Gertner, one of the rare jurists willing to take *Daubert* seriously, agreed,³⁶³ writing that “the NAS Report’s concerns will not be fully met until advocacy changes.”³⁶⁴

A 2009 study of the cases of 137 convicts exonerated by DNA profiling revealed that “[d]efense counsel rarely made any objections to the invalid forensic science testimony in these trials and rarely effectively cross-examined forensic analysts who provided invalid science testimony.”³⁶⁵ One commentator summed it up this way:

³⁵⁸ *Id.* at 4.

³⁵⁹ *Id.* at 7.

³⁶⁰ *Daubert*, 509 U.S. at 597 (“a gatekeeping role for the judge”).

³⁶¹ *Id.* at 596 (citing *Rock v. Arkansas*, 483 U.S. 44, 61 (1987)).

³⁶² D. Michael Risinger, *The NAS/NRC Report on Forensic Science: A Path Forward Fraught with Pitfalls*, 2010 UTAH L. REV. 225, 242.

³⁶³ Judge Gertner wrote seminal cases in several forensic fields. *See* *United States v. Hines*, 55 F. Supp. 2d 62, 67 (D. Mass. 1999) (limiting handwriting testimony); *United States v. Green*, 405 F. Supp. 2d 104 (D. Mass. 2005) (limiting firearms identification testimony); *United States v. Hebshie*, 754 F. Supp. 2d 89 (D. Mass. 2010) (finding ineffective assistance of counsel in an arson case).

³⁶⁴ Gertner, *supra* note 31, at 790.

³⁶⁵ Garrett & Neufeld, *supra* note 87, at 89.

Unlike the extremely well-litigated civil challenges, the criminal defendant's challenge is usually perfunctory. Even when the most vulnerable forensic sciences — hair microscopy, bite marks, and handwriting — are attacked, the courts routinely affirm admissibility citing earlier decisions rather than facts established at a hearing. Defense lawyers generally fail to build a challenge with appropriate witnesses and new data. Thus, even if inclined to mount a *Daubert* challenge, they lack the requisite knowledge and skills, as well as the funds, to succeed.³⁶⁶

In sum, the courts were not solely responsible for *Daubert's* failure. Defense attorneys also bear responsibility. There are limits, however, to what can be expected of overburdened and chronically underfunded public defenders when dealing with expert testimony. Better training for defense counsel (which is sorely needed) is not sufficient. Similarly, access to defense experts (also sorely needed) will not be adequate.³⁶⁷ Defense experts can challenge prosecution experts' methods and opinions but do not have the funds to conduct foundational research, nor to act as an independent evaluator of foundational research on an ongoing basis. The justice system is incapable of providing this expertise. An alternative paradigm is needed.

An independent scientific review is required. NAS has published the most authoritative and independent reviews of forensic science. In addition to the forensic report, NAS has issued reports on sound spectrometry ("voiceprints"),³⁶⁸ two DNA reports,³⁶⁹ polygraph testing,³⁷⁰ and bullet lead analysis.³⁷¹ But NAS is not a governmental entity, and its work depends on funding. The justice system needs scientific expertise on a continuing basis — and thus institutionalized.

The National Commission's proposal, endorsed by PCAST, tasked NIST

³⁶⁶ Neufeld, *supra* note 20, at S110.

³⁶⁷ See Paul C. Giannelli, *Ake v. Oklahoma: The Right to Expert Assistance in a Post-Daubert, Post-DNA World*, 89 CORNELL L. REV. 1305 (2004) (discussing the legal disputes over the scope of the *Ake* — e.g., whether it applied to non capital cases and to non psychiatric experts).

³⁶⁸ NATIONAL RESEARCH COUNCIL, NATIONAL ACADEMY OF SCIENCE, ON THE THEORY AND PRACTICE OF VOICE IDENTIFICATION (1979).

³⁶⁹ NATIONAL RESEARCH COUNCIL, NATIONAL ACADEMY OF SCIENCE, THE EVALUATION OF FORENSIC DNA EVIDENCE (1996); NATIONAL RESEARCH COUNCIL, NATIONAL ACADEMY OF SCIENCE, NATIONAL ACADEMY OF SCIENCES, DNA TECHNOLOGY IN FORENSIC SCIENCE (1992).

³⁷⁰ NATIONAL RESEARCH COUNCIL, NATIONAL ACADEMY OF SCIENCES, THE POLYGRAPH AND LIE DETECTION (2003).

³⁷¹ NATIONAL RESEARCH COUNCIL, NATIONAL ACADEMY OF SCIENCES, FORENSIC ANALYSIS: WEIGHING BULLET LEAD EVIDENCE (2004).

with the responsibility of evaluating forensic disciplines on an ongoing basis.³⁷² It should be adopted. NIST has the expertise and independence for this task and has been increasingly involved in forensic research. There would be a cost, but litigating validity issues across the country at *Daubert* and *Frye* hearings also has a cost. Moreover, there is a significant expense associated with rectifying past mistakes that occurred with hair,³⁷³ bullet lead,³⁷⁴ DNA,³⁷⁵ and arson cases.

Unfortunately, the Attorney General did not even renew the commission's charter in April 2017.³⁷⁶ Instead, he appointed a forensic science working group within the DOJ — headed by a prosecutor instead of a scientist.³⁷⁷ The independent scientists on the commission objected to this action, writing:

The Justice Department now proposes to improve forensic science by moving its oversight and development to an office within the department. This is precisely the opposite of what was recommended by the National Academy of Sciences report and the NCFSS. It is a step backwards, because it reinforces the conditions that contributed to the current problems, namely, placing this discipline within the control of law enforcement and prosecutors. The Justice Department is home to many dedicated public servants including scientists whose passion for justice is unquestioned. However, DOJ is not a scientific body, and it is difficult to see how forensic science can become a true science in that environment. Science flourishes when free and independent; only then can the tools and

³⁷² In 2005, Peter Neufeld proposed an institute of forensic science. Neufeld, *supra* note 20, at S113.

³⁷³ See David R. Cameron, *Forum: Review of FBI Lab Suggests Huge Number of Wrongful Convictions*, NEW HAVEN REGISTER, April 26, 2015 (“The FBI review has identified roughly 2,500 cases that fit those criteria. The review is still in its early stages; thus far, it has considered 268 trials involving 284 defendants. It has found that lab examiners gave flawed testimony regarding the comparison of hairs in 257 of the 268 trials — more than 95 percent. Almost all of the examiners over that period — 26 of 28 — presented flawed testimony.”).

³⁷⁴ See *supra* note 254.

³⁷⁵ See Spencer S. Hsu, *FBI Notifies Crime Labs of Errors Used in DNA Match Calculations since 1999*, WASH. POST, May 29, 2015 (“The FBI has notified crime labs across the country that it has discovered errors in data used by forensic scientists in thousands of cases to calculate the chances that DNA found at a crime scene matches a particular person, several people familiar with the issue said.”).

³⁷⁶ See Spencer S. Hsu, *Sessions Orders Justice Dept. to End Forensic Science Commission, Suspend Review Policy*, WASH. POST, April 10, 2017.

³⁷⁷ See Pem Levy, *Sessions' New Forensic Science Adviser Has a History of Opposing Pro-Science Reforms*, Mother Jones, Aug. 10, 2017 (“But Attorney General Jeff Sessions has resisted efforts to rein in forensic science and hold it to higher standards. And this week, he appointed a senior adviser on forensics who has a history of opposing reforms that would bring more accountability and scientific rigor to forensic crime labs and expert testimony.”).

technology that it creates be truly reliable.³⁷⁸

The American Association for the Advancement of Science concurred, also stressing that independence “cannot be overstated.”³⁷⁹

³⁷⁸ Sunita Sah et al., Observations, *We Must Strengthen the "Science" in Forensic Science*, SCI. AM., May 8, 2017.

³⁷⁹ Spencer S. Hsu, *Science Organizations Renew Call for Independent U.S. Committee on Forensics*, WASH. POST, June 29, 2017 (“The association linked the problem to what it described as an inherent conflict of interest in having law enforcement overseeing the work of forensic labs on which police and prosecutors rely to win and defend convictions.”)

The (Near) Irrelevance of *Daubert* to Criminal Justice and Some Suggestions for Reform

Peter J. Neufeld, JD

In the early morning hours of March 20, 1987, L.T., an 8-year-old girl, slept in her bedroom in the family's Billings, MT, home. An intruder pushed a swing set up against the rear of the house, climbed up, and entered the house through an open second floor bathroom window. Once inside, he quietly entered the little girl's room and raped her. After the assailant fled, L.T. woke her mom and dad, who then summoned the police. The police processed the crime scene, interviewed the victim, and collected the underpants stained with semen and a bed sheet covered with scattered hairs. L.T. gave a description of the perpetrator and worked with a police artist to create a composite sketch. An officer at the police station noticed the sketch and remarked to the case detective that the subject bore a resemblance to Jimmy Ray Bromgard, a 12th grader he had recently arrested for assault after a fight at the local high school. Bromgard volunteered to participate in a corporeal lineup and the girl identified him as the rapist.

The police collected reference blood and hair samples from L.T. and Bromgard to compare with the semen stains on the underwear and the hairs recovered from the sheet. As was often the case before the availability of forensic DNA typing, the attempt to identify the less-sensitive ABO blood type of the semen stain was unsuccessful. Therefore, corroboration of the victim's eyewitness identification depended solely on the microscopic comparison of hairs recovered from the bedding. Hairs from the bed sheet along with reference samples from L.T. and Bromgard were delivered to Arnold Melnikoff, the state's hair expert and Director of the Montana Department of Justice, Forensic Science Division.

Melnikoff was a charter member of the American Society of Crime Lab Directors (ASCLAD); held a master's degree in chemistry; and in an era of forensic generalists, his areas of "qualified" expertise included hair microscopy, toxicology, controlled substances,

Daubert v Merrell Dow Pharmaceuticals, Inc should have an extraordinary impact on criminal litigation, because there is rarely a criminal trial that does not rely on some form of expert testimony. In fact, it is almost irrelevant. Despite the frequency of prosecution proffered scientific and expert testimony in criminal cases, *Daubert* is rarely invoked to challenge it.

In civil cases, when expert testimony is challenged in criminal proceedings, the outcome could not be more different. Because most violent crimes are committed by the poor, their court appointed advocates—overworked and underfinanced—are not up to the challenge. In the absence of a system of effective representation, *Daubert* will not improve scientific evidence in criminal cases. The only way to guard against the misapplication of forensic science is to impose controls and reforms long before the cases come to court. (*Am J Public Health*. 2005;95: S107–S113. doi:10.2105/AJPH.2004.056333)

and arson. Melnikoff issued a report before Bromgard's trial in which he concluded that both a questioned head hair and a pubic hair collected from the victim's bedding had the same microscopic characteristics as the head and pubic hairs collected from Bromgard. At the criminal trial, Melnikoff testified that there was only a 1-in-10,000 chance that the hairs came from anyone other than Bromgard. In fact, there was no scientific basis for Melnikoff's statistical assertion.

No formal challenge to the admissibility of the statistics or to the "match" was ever made. Bromgard was convicted and sentenced to a lengthy prison sentence. On appeal, the Supreme Court of Montana affirmed the conviction citing the persuasive testimony of Melnikoff as overwhelming evidence of guilt.

Eventually, Bromgard contacted The Innocence Project (IP), where I am a codirector. We secured postconviction DNA testing on the semen-stained underpants; the test results excluded Bromgard as the source of the semen. In October 2002, his conviction was vacated and the indictment dismissed. Jimmy Ray Bromgard served 15 years for a crime he did not commit. Moreover, the perpetrator of the rape was not identified and brought to justice.

After Bromgard's exoneration, the IP requested that the questioned head and pubic hairs be microscopically re-examined by the Federal Bureau of Investigation (FBI). They

concluded that Bromgard was not the source of either hair; and that the head hair was microscopically similar to that of the victim.

A Lexis search identified two other cases in which the Montana Supreme Court viewed Melnikoff's hair evidence favorably and affirmed convictions. In both cases, postconviction DNA testing on the original semen samples excluded the convicted offenders, and the men were exonerated.

A NATIONAL CRISIS IN FORENSIC SCIENCE

Forensic science can fail in two ways: (1) lacking reliability (i.e., the inability to reproduce valid results); and (2) bias, incompetence, or a lack of adequate internal controls for the evidence introduced by the forensic scientists and their laboratories. Examples of several problematic "sciences" or failed applications used routinely in forensics illustrate these points.

Hair Microscopy

In a recent FBI scientific paper entitled "Correlation of Microscopic and Mitochondrial DNA Hair Comparisons" the authors found that even the most competent hair examiners make significant errors.¹ In 11% of the cases in which the hair examiners declared two hairs to be "similar," DNA testing

revealed that the hairs did not match. To date, 158 innocents have been exonerated in the United States using postconviction DNA testing. In more than one-third of the cases, the misapplication of forensic science (other than DNA evidence) played a role in convicting the individual; at least 30 of the wrongful convictions relied, in part, on hair “matches.”²

In some jurisdictions, hair microscopy is being phased out and replaced by the more sensitive and discriminating mitochondrial DNA typing test. Many local prosecutors continue to rely on the microscope; mitochondria DNA typing remains relatively expensive and is offered in only a few laboratories. In Alabama, for example, a local prosecutor explained that because of resource constraints, he would continue to rely on the same hair microscopy expert who had previously given erroneous testimony. This expert had testified that hair recovered from a rape victim’s pubic combings were not the victim’s but were consistent with the defendant’s hair. Later, DNA testing excluded the defendant and matched the hair to the victim.

Serology

Serology tests can be reliable, yet in 40% of the DNA exoneration cases, conventional serology had been used by the prosecutor to secure a conviction. The case transcripts reveal that in the vast majority of these cases, the crime lab serologist misrepresented the data to the advantage of the prosecution. The very first postconviction DNA exoneration is illustrative.

At trial, the state serologist testified that the semen found in the mixture of body fluids from the victim’s underpants matched the blood type of the accused, Gary Dotson. What the serologist failed to disclose was that the type also matched the victim’s blood type and that her own fluids could explain the results without implicating Dotson. (Gary Dotson was exonerated by the Cook County, Illinois Criminal Court on August 14, 1989.)

Scientific standards generally prohibit serologists from drawing inferences about the source of the semen when all of the observed genetic markers are consistent with the victim. In dozens of IP cases, however, state crime laboratory serologists ignored these

rules and engaged in prosecution bias. The misconduct involves at least 20 serologists in as many different states.

Fingerprinting

Wrongful convictions have also resulted from the misapplication of fingerprint evidence, yet law enforcement, especially the FBI, stubbornly resisted any challenge to this forensic sacred cow. Stephen Cowans was convicted in 1997 of shooting a Boston police officer. Two fingerprint experts told a jury during the trial that a thumbprint left by the perpetrator was “unique and identical” to Cowans’ print because it matched at 16 points. In 2004, post-conviction DNA testing on several items of evidences excluded Cowans as the perpetrator.³

A new prosecutor assigned to the case wanted to resolve the apparent contradiction between the two most discriminating forensic sciences. The prosecutor had the thumbprint re-examined by state (rather than city) police experts who quickly concluded that Cowans was not the source of the print. (In the aftermath of the Cowans case, the Boston Police Department retained an outside auditor to review the case. The audit resulted in the department shutting down the latent fingerprint unit.)

In Spring 2004, terrorists bombed the railroad in Madrid, killing nearly 200 people. The Spanish National Police recovered a plastic bag of detonator caps inside a van parked close to the bombsite that were similar to those used in the railroad bombings. At least one sufficiently detailed latent fingerprint was observed on the plastic bag. A digital image of the print was forwarded to the FBI and run through the Bureau’s Automated Fingerprint Identification System. The computer selected more than a dozen potential “matches,” and a senior FBI fingerprint examiner compared the digital print visually to the numerous “hits.” The fingerprint was matched to a Portland, OR resident, named Brandon Mayfield, who had converted to Islam, regularly attended a mosque, and married an Arab woman. Mayfield’s fingerprints were contained in the Automated Fingerprint Identification System, database because he had served in the US armed forces. Two FBI fingerprint experts swore in affidavits that they

were 100% certain that the prints belonged to Mayfield. When the Spanish police ultimately arrested the real source of the fingerprint, the FBI initially defended their mistake as the result of a poor digital image.⁴ But if it was so poor, how could they have been 100% certain? And why would the FBI rely routinely on imprecise digital images to provide support for local police departments?

Compositional Analysis of Bullet Lead

Within a week of the Cowans’ exoneration, the National Research Council delivered a stinging critique of another forensic method the FBI has relied upon for more than 20 years.⁵ The technique, Compositional Analysis of Bullet Lead (CABL), compares the quantity of various elements that comprise a lead slug recovered from a crime scene with the composition of the lead found in unused bullets seized from a suspect. In criminal cases, to say that two samples match or are similar is potentially relevant, but, unless the jury learns just how rare the match is, its probative value is minimal.

In Compositional Analysis of Bullet Lead cases where the FBI lab described the two bullet samples as indistinguishable, FBI analysts would routinely testify that the bullet from the body came from the same box of ammunition as those found in the defendant’s possession. The problem, according to the National Research Council, is that there simply is no scientific basis for such a conclusion. (The FBI deserves credit for requesting the NRC study.)

The FBI never subjected its validation process for peer review or any meaningful internal review of the rigor of its results. Regrettably, the same criticisms can be directed at almost every forensic discipline which attempts to “match” or individualize crime scene evidence. The only difference is that the other forensic techniques have not been scrutinized by the National Research Council.

DNA Typing

In the past 10 years, state and local crime laboratories have been the focus of ad-hoc probes because of flaws and misconduct exposed by postconviction DNA exonerations. In contrast to most forensics, DNA testing has been carefully examined by the National

Research Council and its methods plainly validated. Yet even DNA typing, perhaps the most rigorous forensic discipline, has not been above the fray. The Virginia Division of Forensic Science, for example, issued a formal analysis in 2000, in the Earl Washington case, claiming that the sperm recovered from the deceased victim's vagina originated from an unidentified man. In reality, the semen originated from a known convicted rapist whose DNA profile had previously been matched to the semen recovered from the blanket where the victim was sexually assaulted and stabbed. Testing performed in 2004 by an independent laboratory proved conclusively that the vaginal sperm, like that found on the blanket, matched the convicted rapist.⁶ Sloppiness in the Virginia crime lab had produced an erroneous result but the supervisors refused to acknowledge error nor take corrective action.

Bad forensic science is bad law enforcement. Each time unreliable science, incompetent scientists or crime lab misconduct is used to arrest, indict, or convict an innocent person, the real perpetrator remains free to commit more crime. Faulty forensic science may wrongly exclude suspects. Guilty defendants can be wrongfully exculpated. In criminal cases that use forensic science during the investigation and trial, meaningful precautions must exist to guard against junk science and unreliable results.

THE UNEVENHANDED APPLICATION OF DAUBERT

Ten years after the United States Supreme Court decided *Daubert*,⁷ many in the judiciary, the legal academy, and the scientific community herald it as one of the most important decisions of the last century. It obligated trial court judges to assume the role of “gatekeepers” and to exclude proffered scientific evidence unless it rested on scientifically valid reasoning and methodology. Many thought *Daubert* would be the meaningful standard that was lacking in criminal cases and that it would serve to protect innocent defendants.

In *Kumho Tire*,⁸ the Court extended the holding of *Daubert* to all expert evidence, even if not based on hard science. But it is

not a coincidence that both of these cases and almost all of the post-*Daubert* federal appellate decisions that further defined the standard have been civil rather than criminal.

In theory, *Daubert* should have an extraordinary impact on criminal litigation. In 25 years, I have not tried a criminal case in which the prosecutor did not offer some form of expert evidence. Yet, despite the frequency with which scientific and expert testimony is proffered in criminal cases, there is a dearth of *Daubert* challenges and hearings. When the issue is raised in criminal proceedings, the outcome is vastly different than what occurs in civil cases.

An analysis of post-*Daubert* decisions demonstrates that whereas civil defendants prevail in their *Daubert* challenges, most of the time criminal defendants almost always lose their challenges to government proffers.⁹ But when the prosecutor challenges a criminal defendant's expert evidence, the evidence is almost always kept out of the trial. This is true in both federal and state courts. And even though *Frye*¹⁰ remains the test in more than a dozen states, criminal defendants fared no better under *Frye*. In the first 7 years after *Daubert*, there were 67 reported federal appellate decisions reviewing defense challenges to prosecution experts. The government prevailed in all but 6, and even among the 6, only 1 resulted in the reversal of a conviction. In contrast, in the 54 cases in which the defense appealed a trial court ruling to exclude the defendant's expert, the defendant lost in 44 cases. In 7 of the remaining 10, the case was remanded for a *Daubert* hearing.

One case is particularly revealing of the inadequacy of *Daubert* in criminal cases. An Oklahoma state hair expert matched 17 hairs found at the murder scene to two defendants, Ron Williamson and Dennis Fritz. A federal district court judge, in what may be the only successful post-*Daubert* challenge to hair microscopy, indicated that the expert's hair-comparison testimony failed to meet any of the requirements of *Daubert*.¹¹ After the conviction was vacated, the 17 hairs were re-examined using mitochondria DNA testing, and none of them matched either defendant. Nevertheless, in the next 7 years, no other court has concluded that expert hair comparison is not sufficiently scientific. In fact, when a

scandal involving another Oklahoma serologist/hair examiner exploded in the press and the State agreed to re-examine microscopically the hair evidence in dozens of closed cases, the same expert it used in the Williamson trial was enlisted by the State for the re-examination, with full knowledge that he had been wrong 17 times in 17 attempts.

The Fallacy of the “Crucible of the Court”

For years in the forensic science community, the dominant argument against regulating experts was that every time a forensic scientist steps into a courtroom, his work is vigorously peer reviewed and scrutinized by opposing counsel. A forensic scientist might occasionally make an error in the crime laboratory, but the crucible of courtroom cross-examination would expose it at trial. This “crucible,” however, turned out to be utterly ineffective.

In not one of the half-dozen most sensational forensic-science scandals of the last 20 years, involving serial fraud and gross misconduct, were the transgressions of “experts” revealed by defense counsel at trial. Dr Ralph Erdman, a Texas state medical examiner, for example, conducted “zipperless” autopsies (without a single incision) and got away with it because the defense failed to seek independent autopsies or re-examinations of postmortem samples. West Virginia's chief serologist, Fred Zain, frequently “dry labbed” serology tests without conducting the actual laboratory experiment in order to produce reports helpful to the police. After a post-conviction DNA exoneration, 35 of Zain's cases were reviewed. In all 35 cases, his bench notes did not support the findings contained in his final reports.¹² His fabrications had not been exposed, because the defendants' lawyers never bothered to review the bench notes.

New York state troopers, in another example, framed more than a dozen defendants with phony fingerprints.¹³ Not until one of the troopers bragged about his exploits at a job interview with the CIA did the truth surface. An investigation by the special prosecutor revealed fabrication in more than three dozen criminal cases over a decade. Five troopers pleaded guilty. Similarly, the Montana State Crime Laboratory Director made up powerful,

albeit false, results for years¹⁴, and the scientifically illiterate Montana judiciary relied on those results to affirm convictions. Not a single defense attorney called in an opposing expert to challenge the data.

In one third of all postconviction DNA exonerations, unexposed scientific fraud, the criminalist's incompetence, or an expert's reckless disregard for the truth at the trial was a significant cause of the wrongful conviction. Regardless of the *Daubert* standard, without zealous investigation and cross-examination of the proffered expert evidence, many improper and even fraudulent uses of scientific data are not exposed.

Statistics substantiate the ubiquity of defense failure to initiate *Daubert* challenges, confirming the rarity in the trial courts of any defense challenge to a prosecutor's proffered expert testimony. State courts receive 200 times more criminal prosecutions than federal courts. Forensic science is used most commonly in crimes of violence, and most crimes of violence are tried in state court. In the first 7 years after *Daubert*, there were a mere 211 reported challenges to prosecution experts in state court, and the prosecution defeated the challenge 161 times. From August 1999 through August 2000, there were only 50 reported challenges to admissibility citing *Daubert* in state criminal cases; in 2000, nearly 15 million criminal filings were made in the State. If no one challenges the speculative science or scientist, there is nothing for a gatekeeper to tend to. Thus, the principal failing of *Daubert* is its misplaced reliance on a robust adversarial system to expose bad science. In reality the playing field is not level, and the system is anything but robust.

Poorly Funded, Unskilled Counsel; an Inadequate Pool of Experts

Why are there so few challenges from criminal defendants' lawyers? Most criminal defendants are indigent. They are represented by public defenders, contract defenders, and private lawyers paid minimal fees by the government. In most states, before an assigned counsel can retain an expert to educate him or her, review the opposing expert's data or conduct independent testing, counsel must secure approval from the presiding judge, an elected county official. The money

to pay for the expert comes from a strained county treasury, and judges are reluctant to authorize expenditures for experts. Unlike prosecutors with free access to government medical examiners and publicly funded crime labs, defense counsel must usually seek independent contractors, and then, if the client is indigent, only with the court's permission. The FBI supplies free services, but only for prosecutors and police.

Unlike the extremely well-litigated civil challenges, the criminal defendant's challenge is usually perfunctory. Even when the most vulnerable forensic sciences—hair microscopy, bite marks, and handwriting—are attacked, the courts routinely affirm admissibility citing earlier decisions rather than facts established at a hearing. Defense lawyers generally fail to build a challenge with appropriate witnesses and new data. Thus, even if inclined to mount a *Daubert* challenge, they lack the requisite knowledge and skills, as well as the funds, to succeed.

Lawyers are not the only problem—judges have to share some responsibility. In *Barefoot v Estelle*,¹⁵ the Supreme Court sustained the admission of psychiatric testimony during the penalty phase of a capital case from Dr James Grigson who, without ever examining Mr. Barefoot, opined under oath that there was “. . . a one hundred percent and absolute chance that Barefoot would commit future acts of criminal violence.” Justice Blackmun, who years later authored *Daubert*, dissented: “In the present state of psychiatric knowledge this is too much for me. One may accept this in a routine lawsuit for money damages, but when a person's life is at stake . . . a requirement of greater reliability should prevail. In a capital case, the specious testimony of a psychiatrist, colored in the eyes of an impressionable jury by the inevitable untouchability of a medical specialist's words, equates with death itself.”

Justice Blackmun's principled critique of the adjudicative process is plainly correct. The reality is that if a corporation is sued for millions of dollars in a toxic tort case, plaintiffs' attorneys hire scientific experts because they stand to share in any settlement or award. The substantial legal fees paid by the corporation enable civil defendants to secure the services of equally well-regarded experts.

Judges consider the science with far greater scrutiny and caution.

If You Can't Discover the Underlying Data, There is Nothing to Challenge

The discovery available by statute and case law to a defendant who is sued for money greatly exceeds the discovery available for a defendant facing execution. In Texas, the state that leads the nation in executions, a criminal defendant is not by statute entitled to see before trial the laboratory bench notes for tests conducted on the case evidence. All that he gets is a conclusory report without the underlying notes. In Virginia, the state that is second to Texas in executions, the state's highest court has explicitly held that a defendant facing execution is prohibited from reviewing the bench notes of the state forensic scientist who will be providing the most inculpatory evidence at trial.¹⁶

In the 1999 rape prosecution of Josiah Sutton, a Houston Police Department Crime Lab DNA report furnished to Sutton's attorney was sparse. It merely stated that the DNA testing of Sutton's semen could not exclude him. Four years after his conviction, and once the notorious Houston Crime Lab scandal broke in the press,¹⁷ the bench notes were disclosed to Sutton's attorney for the first time. The notes revealed that in all likelihood Sutton was excluded as a source of the semen in the original testing. Additional DNA testing in 2003 confirmed Sutton's innocence in the case, and he walked out of prison.¹⁸

Beware of Experiments Conducted for Purposes of Litigation

Clinical laboratories deserve greater public confidence than crime laboratories because of the relatively extravagant validation studies and rigorous review by the Food and Drug Administration that precede the transfer of laboratory technology from research to clinical application. With the exception of DNA analysis and possibly a few other disciplines, no validation studies or “trials” exist for most forensic sciences before the technology is used in real casework. Professional fingerprint publications, unlike scientific publications, do not contain critiques and reanalysis by other scientists, but instead focus on how to lift prints.

When *Daubert* was remanded to the 9th Circuit, Judge Kozinski warned of the dangers of giving too much credence to scientific tests conducted for purposes of litigation. Whereas he favored greater deference to findings derived from academic research, he cautioned that the objectivity of the scientist and, thus, the results generated were compromised whenever the experiment is conducted for a specific case.¹⁹ In writing his *Daubert II* opinion, Judge Kozinski eliminated criminal litigation from his caution about science conducted for litigation purposes. He reasoned that in criminal cases, *all* scientific experiments are conducted for the purpose of litigation. Instead of realizing the inherent danger of bias in the experiments conducted during criminal investigations, Judge Kozinski, without proper justification, dismissed this concern altogether.

The danger is neither abstract nor hypothetical. The “examiner bias” phenomena is well known in most applied sciences. In clinical medicine, for example, studies have documented unintended bias resulting from the examiner’s exposure to irrelevant case information, increasing the likelihood of a false positive.²⁰ In criminal investigations, it is routine for police to offer a detailed narrative of the crime and an inventory of whatever other inculpatory evidence they have against the suspect on the request form used to order a particular scientific test. A 1997 Department of Justice Inspector General’s investigation of the FBI revealed that examiners in some units knew of the conclusions of examiners in other units and tailored their own conclusions to be consistent. (The Stacey Report found that examiner bias—and not the quality of the digital image—was the primary cause of the false match in the Madrid train bombing case. Moreover, the panel found that independent verification is likely to fail and examiner bias is of greater danger in high profile cases.)

The truth is most crime labs work hand-in-glove with law enforcement agencies and prosecution services. Examiner bias is systemic. Prosecution “team” identification is chronic. Thus, even the logical constructs imposed by *Daubert* on trial courts are not evenly applied to the forensic and clinical sciences.

REFORMS UPSTREAM OF THE COURTHOUSE

The IP’s detailed examination of police and expert investigations, prosecutions, and trials of the first 138 of 158 postconviction DNA exonerations provides unprecedented insight into the most prevalent causes of wrongful convictions. Because of the conclusive, objective nature of DNA exonerations, these cases offer a unique window into systemic flaws that lead to unjust results. In more than 33% of the DNA exonerations, material misstatements of fact by forensic scientists played a significant role in the wrongful convictions.² Most common were (1) conclusions without any scientific basis; (2) reports that ignore data or deliberately distort data; (3) testimony that ignores or deliberately distorts the report; and (4) testimony and report writing clearly beyond the competence of the examiner. None of these material misstatements of fact were adequately revealed and remedied during the trial or appellate process, and but for the fortuitous DNA exoneration, they would have remained obscured.

The judicial process has failed to provide obligatory controls to ensure the fairness of the proceeding. If the courts cannot be relied on for this protection, other remedies must be found further “upstream” so that the disreputable evidence is never proffered. Below I describe possible reforms: systems of audits, accreditation and quality assurance, and validation. The suggested reforms start from the premise that although greater judicial oversight would be welcome, in the absence of a meaningful, well-funded, and well-staffed indigent criminal defense system, the courts will be unable to provide effective relief no matter how rigorous the standard for admissibility of scientific evidence. The sequence of reforms, similarly, is prioritized according to which is most doable in the shortest time.

Reform 1: Independent External Audits to Investigate Instances of Misconduct or Gross Negligence

To enhance the integrity of forensic science results, Congress might encourage a simple yet fundamental program: external independent audit and investigation. Congress has provided generous support for forensic DNA

typing, but experts estimate that only 20% of violent crime investigations will benefit from evidence suitable for DNA testing. Because other forensic disciplines lack the heightened scientific dimension of DNA, measures are needed to raise their standards of performance. There is simply no better way than external audits to investigate the scope of a problem and to remediate, thus reducing the risk of it happening again.

The US Department of Justice Inspector General delivers independent oversight to the FBI crime laboratory. When it was revealed that a FBI crime lab scientist failed to follow a required control in casework, instead of relying on the Bureau’s internal affairs mechanism, the Inspector General (IG) opened an independent investigation to assess the scope of the failure, the potential impact on prosecutions, the reason existing quality controls failed, and to recommend remedial action to reduce the risk of recurrence.²¹ Although the FBI lab is accredited by the American Society of Crime Laboratory Directors—Lab Accreditation Board (ASCLD/LAB), the routine internal audits and external inspections currently mandated by ASCLD/LAB do not (nor are they meant to) address these special circumstances. As a consequence of the IG investigation, the Bureau retested evidence in more than 100 cases assigned to the reckless scientist. The reality in most states is strikingly different. Generally, the states have no established and readily accessible independent forensic auditors. All too frequently, local officials with a stake in the outcome of the investigation refuse to give up control of the inquiry to independent experts. In Montana, in the aftermath of the Bromgard exoneration, a peer-review committee comprised of the nation’s top hair examiners urged the Montana Attorney General to establish an external independent audit committee to re-examine the hairs in the few hundred felony cases in which Melnikoff provided hair analysis. Montana’s Attorney General, who had supervisory responsibility for several of the prosecutions that relied on Melnikoff and who personally relied on Melnikoff when he was a local prosecutor, refused to appoint an independent investigator and refused to order the re-examination of Melnikoff’s Montana casework. The ongoing scandal at the Houston

Police Department Crime Laboratory reveals that erroneous results were produced in several cases and in more than one forensic discipline. But the investigation had been obstructed by a local county prosecutor who, as a matter of routine, relied on the lab in thousands of criminal prosecutions. Despite the protest of elected state and local officials demanding that the district attorney's conflict of interest mandates his recusal, he refused to step down. Ultimately, responding to a tidal wave of bad press, the city government contracted with an independent auditor with extensive expertise.

Congress might require independent external investigations into allegations of serious negligence or misconduct committed by employees or contractors of the forensic laboratory, as a condition of federal funding to state and local crime labs. Ultimately, the audit function should illuminate what went wrong and how to make it right, thereby reducing the risk of future mishaps. The essential elements of the certification would include:

- (1) Investigators must be independent of the entity being investigated. Investigators do not report to or depend on the laboratory for any resource or benefits. Investigators do not rely on the results of the laboratory in a professional capacity.
- (2) Investigators must have adequate experience and qualifications and be trained in conducting similar reviews.
- (3) Resources must be adequate to conduct a professional and thorough investigation.
- (4) Protocols must be established for conducting investigations.
- (5) Adequate quality control for the investigation must be established.

A standard format for the report must be established, with a presumption that the report will be made public. The report shall address the individual conduct and scope and, where appropriate, make systemic recommendations for improvement and order re-examination of casework. The report should contain adequate documentation and support for the findings.

It makes more sense to delegate this investigatory responsibility to the states than centralize it with the Department of Justice IG. The expansive and extensive oversight neces-

sary to monitor adequately all forensic disciplines in all state and local laboratories receiving federal funding could overwhelm the IG. The load is lightened considerably if spread among the states. Moreover, there is broad concern that state criminal justice systems should have the freedom and flexibility to implement their own integrity controls. A concern for federalism can be satisfied if Congress delegates to the states the responsibility of creating or identifying a pre-existing independent investigative mechanism but at the same time requires that the state system be certified by the Department of Justice IG. (One week before the 2004 election, the President signed into law the Justice For All Act (Public Law 108-711). Section 311(4) of the law provides that as a condition of receiving Coverdell federal grant money to aid state and local crime labs, states are required to certify that ". . . a government entity exists and an appropriate process is in place to conduct independent external investigations into allegations of serious negligence or misconduct substantially affecting the integrity of the forensic results . . .")²²

Reform 2: A National System of Accreditation and Quality Assurance, and Independence from Law Enforcement

Forensic science is to criminal justice what clinical laboratory science is to health care. Health and public safety depend on the integrity of the product. The consumer of clinical medicine receives a measure of protection through government-imposed and -regulated quality assurance and quality control. Defendants, victims, and the public would derive comparable protection from government-imposed oversight to ensure the integrity of forensic science before it gets to court. But whereas a national regulatory scheme has been in place for clinical laboratories since 1968, there is simply no national or, with one exception, meaningful state regulation of forensic science. Instead, the protections to avoid compromised evidence are few, and the measures to investigate and address abuses once they are discovered are virtually nonexistent.

There is an excellent model for the regulation of government crime laboratories: The Clinical Laboratories Improvement Act of

1967, amended in 1988. The Act established a system of accreditation and proficiency testing for clinical laboratories that service the medical profession. In contrast, with the exception of New York, no state or local crime laboratory is actively regulated by any government agency. The problem is exacerbated by the lack of any formally enforced objective criteria for interpreting, reporting, and testifying about forensic data.

Most of the crime laboratories are resistant to any oversight. Additionally, in an effort to fend off a Clinical Laboratories Improvement Act-type regulatory approach, some public crime lab directors have urged their colleagues to voluntarily seek accreditation through their private professional organization, ASCLD/LAB. Paul Ferrara, chief of Virginia's Bureau of Forensic Science, told Congress that just as doctors and lawyers regulate themselves through professional associations, so should crime labs. This is simply not true. (In part, as a result of laboratory's errors in the Earl Washington case, Virginia enacted legislation that, for the first time, will provide some measure of scientific oversight of the state's crime labs.)

Although unquestionably, ASCLD/LAB fulfills a critical role in the overall improvement of the delivery of forensic services, they cannot be the final arbiter. When lawyers or doctors engage in misconduct, there are government institutions that intervene. When matters of public health and safety are at stake, the American Medical Association and the American Bar Association do not have the last word. How would consumers react to selecting meat at the grocery if the label read "Certified by the Meatpackers Association," instead of the US Department of Agriculture?

Many of the forensic abuses are indicative of prosecution bias. These laboratories should be independent of police control. The government, not private laboratories, produce almost all of the forensic evidence offered by prosecutors. Of the government labs, 80% are controlled by police; most, if not all, will only examine evidence submitted by the police or a district attorney. The Josiah Sutton exoneration precipitated a broader review of the Houston Police Department Crime Laboratory. One of the more serious lapses revealed in a superficial investigation is that the

underlying data do not support the conclusions of the forensic scientist in the official report or in courtroom testimony. In both DNA and ballistics cases, Houston police criminalists misrepresented the data to advance the prosecution theory of guilt. This can be prevented, in part, by “blind” testing and prohibiting examiners from receiving crime data extraneous to the specific scientific test. Laboratories need to control the flow of information from police to the forensic scientist. They can continue to assist law enforcement and prosecutors without performing as subordinates. In some jurisdictions, the office of medical examiner serves this purpose. But unfortunately, all too frequently, the medical examiner also sees itself as a member of the prosecution team.

Reform 3: A National Institute to Validate Technologies, Methodologies, and Set Standards for Interpretation of Data

Basic research for medicine is underwritten by the National Institutes of Health. Both Clinical Laboratories Improvement Act and the Food and Drug Administration provide essential controls for the interpretation of clinical laboratory data. There is nothing comparable for forensic science. Truly independent forensic research does not exist. Most of the studies are commissioned by the Department of Justice and carried out by the crime labs with a significant bias in the outcome. For most forensic science, there are no enforceable standards for individual interpretation of data. The President’s DNA and Forensic Science initiative announced during the summer of 2003 acknowledges the seriousness of the present deficiencies and calls for the establishment of a National Forensic Science Commission to spearhead the effort to improve the delivery of forensic services. That commission is included in the Justice For All Act.

These problems could be remedied by the creation of an institute of forensic science, jointly operated by a medical school and law school or as a necessary extension of the National Academies of Science. It could provide the necessary conflict-free environment augmented by rigorous academic policies and procedures. Federal grants to the institute could finance objective research, necessary

validation studies, and peer review. Moreover, the synergy of law and medicine would enhance the development and implementation of appropriate standards and controls for reporting scientific results in writing and in court. The effort will fail, however, unless it is managed jointly by scientists and legal scholars who are independent of as well as those who work with law enforcement.

CONCLUSIONS

Although scientific evidence is often more reliable than other types of evidence, not all that purports to be “science,” is. Rules of admissibility promulgated by courts and legislatures do not function well in a criminal justice system devoid of effective defense for indigent defendants. Thus, intervention and quality control must occur further upstream in the process. The easiest quality control would be to institutionalize external investigations. But more is needed. Government oversight and the creation of independent academic centers to validate technologies and techniques, encourage best practices, and enforce appropriately cautious standards for the interpretation of data could dramatically enhance the reliability of forensic science and engender greater public confidence in the outcome. ■

About the Author

The author is with the Innocence Project at the Benjamin N. Cardozo School of Law, New York, NY.

Request for reprints should be sent to Peter J. Neufeld, J.D., *The Innocence Project*, 100 Fifth Avenue, New York, N.Y. 10011 (e-mail: info@innocenceproject.org).

This article was accepted December 27, 2004.

Acknowledgments

This article was written with the assistance of Innocence Project policy analyst Sarah Tofte.

References

- Houck M, Budowle B. Correlation of microscopic and mitochondrial DNA hair comparisons *J Forensic Sci*. 2002;47(5). Available at: <http://journalsip.astm.org/journals/forensic/toc02/4752002.htm>. Accessed June 10, 2005.
- Unpublished research conducted by the Innocence Project.
- Smalley S. “Police Shutter Print Unit.” *The Boston Globe*, October 14, 2004. In the aftermath of the Cowans case, the Boston Police Department retained

an outside auditor to review the case. The audit resulted in the department shutting down the latent fingerprint unit.

- The Inspector General of the Department of Justice has launched an investigation. The FBI conducted its own review, recruiting a panel of independent experts. The panel concluded that the problem was not the quality of the digital image but rather the unprincipled bias of the FBI examiners. Remedial action was taken. See the FBI report by Robert B. Stacey, entitled “Report on the Erroneous Fingerprint Individualization in the Madrid Train Bombing Case,” Available at: http://www.fbi.gov/hq/lab/fsc/current/special_report/2005_special_report.htm. Accessed December 22, 2004.
- “Forensic Analysis: Weighing Bullet Lead Evidence,” National Research Council, 2004.
- Green F. Subsequent work points to another suspect in the slaying, but a state doctor is unmoved. *Richmond Times Dispatch*, June 5, 2004.
- Daubert v Merrell Dow Pharmaceuticals, Inc*, 509 US 579 (1993).
- Kumho Tire Co v Carmichael*, 526 US 137 (1999).
- Risinger DM. Navigating expert reliability: are criminal standards of certainty being left on the dock? 64 *Albany Law Review*, 2000.
- Frye v United States*, 54 App. DC 46, 293 F. 1013.
- Williamson v Reynolds*, 904 F.Supp 1529 E.D. Okla (1995).
- In the Matter of an Investigation of the West Virginia State Police Crime Laboratory, serology Division* (WVa 1993) 438 S.E.2d 501 (Zaine I); and 445 S.E.2d 165 (Zain II).
- Nordheimer J. “Trooper’s fall shakes both police and public.” *New York Times*. November 15, 1992:41.
- Possley M. “Lab errors cited in petition on 200 Montana cases.” *Chicago Tribune*. August 26 2004;12.
- Barefoot v Estelle*, 463 US 880, 103 S. Ct. 3383 (1983).
- Spencer v Virginia*, 384 S.E.2d 785, 791 (Va 1989)
- Kimberly J. “Reaction to Scandal: House Passes Crime Lab Bill.” *Houston Chronicle*. May 2, 2003:A30.
- Khanna R. “Rest of DNA May Free Convict.” *Houston Chronicle*. March 11, 2003:A1.
- Daubert v Merrell Dow Pharmaceuticals, Inc*, (Daubert II), 43 F.3D 1311 (1995).
- Saks MJ, Risinger DM, Rosenthal R, Thompson WC. Context effects in forensic science: a review and application of the science of science to crime laboratory practices in the United States. 43 *Science & Justice* 77-90 (2003).
- The FBI DNA laboratory report: a review of protocol and practice vulnerabilities. Federal Bureau of Investigation (news release), May 27, 2004. A copy of the IG report can be found at www.usdoj.gov/oig/igwhnew1.htm.
- 42 USC 3797 k(4).

112 Nw. U. L. Rev. Online 1

Northwestern University Law Review Online

September 12, 2017

Jennifer D. Oliva, Valena E. Beety^{a1}

Copyright © 2017 by Jennifer D. Oliva and Valena E. Beety

DISCOVERING FORENSIC FRAUD

ABSTRACT--This Essay posits that certain structural dynamics, which dominate criminal proceedings, significantly contribute to the admissibility of faulty forensic science in criminal trials. The authors believe that these dynamics are more insidious than questionable individual prosecutorial or judicial behavior in this context. Not only are judges likely to be former prosecutors, prosecutors are “repeat players” in criminal litigation and, as such, routinely support reduced pretrial protections for defendants. Therefore, we argue that the significant discrepancies between the civil and criminal pretrial discovery and disclosure rules warrant additional scrutiny.

In the criminal system, the near absence of any pretrial discovery means the criminal defendant has little to no realistic opportunity to challenge forensic **evidence** prior to the eve of trial. We identify the impact of pretrial disclosure by exploring the admission of expert **evidence** in criminal cases from a particular forensic discipline, specifically forensic odontology. Finally, this Essay proposes the adoption of pretrial civil discovery and disclosure rules in criminal proceedings to halt the flood of faulty forensic **evidence** routinely admitted against defendants in criminal prosecutions.

INTRODUCTION	2
I. BACKGROUND	5
II. PRETRIAL RULES FAVOR THE REPEAT LITIGANT	7
III. PRETRIAL DISCOVERY IN CRIMINAL AND CIVIL PROCEEDINGS	9
IV. LACK OF PRETRIAL DISCOVERY IN CRIMINAL PROCEEDINGS CONTRIBUTES TO THE CONTINUED ADMISSION OF FAULTY FORENSICS	11
V. SOLUTION: PRETRIAL DISCOVERY AND DISCLOSURE IN CRIMINAL PROCEEDINGS	13
CONCLUSION	17

2 There is no justification for accepting that a method is valid and reliable in the absence of appropriate empirical **evidence Forensic science is at a crossroads.*^{d1}

INTRODUCTION

On June 22, 2017, the Supreme Court decided *Turner v. United States*.¹ The question before the Court concerned the scope of the prosecutorial *Brady*² obligation to disclose to the defense **evidence** favorable to criminal defendants. The *Turner* defendants, who were convicted of the brutal 1984 robbery and murder of a middle-aged mother of six, steadfastly litigated their innocence. The crux of their argument before the Supreme Court was that prosecutors had suppressed witness statements about a possible alternative perpetrator in violation of *Brady*.

The federal government did not deny that it had failed to turn over **evidence** favorable to the defense pretrial. Instead, it relied exclusively on the technical argument that the at-issue alternative suspect statements were “immaterial.” The Supreme Court agreed and held that the state's suppression of the witness statements did not run afoul of *Brady*. The *Turner* case shines a harsh light on criminal defendants' extremely limited right to pretrial discovery. Moreover, and as demonstrated by the 2016 President's Council of Advisors on Science and Technology (PCAST) Report, *Forensic Science in Criminal Courts*,³ this lack of robust pretrial discovery can result in the admission of **unreliable** scientific **evidence** and, ultimately, wrongful convictions in criminal proceedings.

*3 We believe that certain structural dynamics that dominate criminal proceedings significantly contribute to the admissibility of faulty forensic science in criminal trials. We also believe that these dynamics are more insidious than questionable individual prosecutorial or judicial behavior. Not only are judges likely to be former prosecutors,⁴ prosecutors are “repeat players” in criminal litigation and, as such, “typically seek to reduce pretrial protections that would impede [their] intentions.”⁵ Therefore, we argue that the significant discrepancies between the civil and criminal pretrial discovery and disclosure rules warrant additional scrutiny.

Legal commentators routinely espouse that the rules of criminal procedure provide trial-based protections to defendants superior to those applicable to any other litigants in the legal system.⁶ Even assuming the truth of that claim, the rules of civil procedure provide many of these protections and concomitant transparency *throughout the pretrial proceedings*, during which the overwhelming majority of cases in both the criminal and civil systems are resolved.⁷ The civil system's unfettered *4 access to pretrial discovery allows both litigants and judges to thoroughly scrutinize the reliability and validity of proffered forensic **evidence** before a case goes to trial and, necessarily, before any party's experts are allowed to testify.⁸ In the criminal system, on the other hand, the near absence of any pretrial discovery means the criminal defendant has little to no realistic opportunity to challenge forensic **evidence** prior to the eve of trial.⁹

The pretrial rules pertaining to prosecutorial disclosure are gradually moving in the direction of increased transparency. But they have not yet evolved either to ensure timely, pretrial disclosure of relevant **evidence** to the defense or to effectively combat the admission of flawed forensic **evidence** repeatedly introduced against defendants in criminal cases. Unlike in civil cases, criminal courts often automatically accept, rather than thoroughly vet, forensic testimony, irrespective of its scientific reliability and validity.¹⁰ Under Georgia law, for example, an opposing party cannot even challenge the ability of an expert to testify in criminal proceedings because the legislature has decreed that such opinions “shall always be admissible.”¹¹ Georgia's civil expert witnesses, by comparison, are subject to the rigorous pretrial vetting rules provided by the Federal Rules of **Evidence** and Civil Procedure.¹² As one commentator explains, the substantial discrepancies between civil and criminal expert **evidence** *5 gatekeeping are “particularly unacceptable given the law's claim that inaccurate criminal convictions are substantially worse than inaccurate civil judgments, reflected in the different applicable standards of proof.”¹³

This Essay examines systems-level procedural problems that all too often contribute to the admission of flawed forensics in criminal proceedings. We begin by examining the concept of the “repeat litigant” and its role in shaping the applicable evidentiary standards in both civil and criminal cases. Next, we highlight the discrepancies between the pretrial discovery and disclosure rules applicable in civil and criminal cases, and how they exacerbate the repeat litigant advantage of prosecutors. We then identify the impact of these variant rules by exploring the admission of forensic odontology, or bite mark, **evidence** in criminal cases. Finally, this Essay proposes the adoption of pretrial civil discovery and disclosure rules

in criminal proceedings to halt the flood of faulty forensic **evidence** routinely admitted against defendants in criminal prosecutions.

I. BACKGROUND

The September 2016 PCAST Report,¹⁴ like the NAS Report before it,¹⁵ challenged forensic disciplines to reform and implored the criminal justice system to stop admitting faulty science to convict innocent people. The PCAST Report recommendations also closely tracked Federal Rule of Evidence 702's expert witness admissibility requirements, expounded upon by the *Daubert* decision, that experts offer some kind of specialized knowledge, that their testimony be based on sufficient facts or data, and *6 that it be the product of reliable methodology that has been properly applied to the present case.¹⁶ Remarkably, the Department of Justice and Federal Bureau of Investigation--that is, the federal prosecutors and police--refused to adopt the PCAST Report recommendations aimed at ensuring that only scientifically valid and reliable **evidence** is admissible in the criminal courtroom.¹⁷

Articles traditionally argue that bad science permeates criminal proceedings for at least three reasons: (1) lawyers (including judges, prosecutors, and defense attorneys) lack scientific aptitude; (2) judges, many of whom are former prosecutors, have a pro-prosecution bias; and (3) prosecutors are more focused on securing convictions than reaching a just result.¹⁸

But we argue that these observations miss a crucial question: Why do judges frequently fail to keep faulty forensics out in criminal cases despite the fact that they rigorously enforce *Daubert's* gatekeeping requirements when presiding over civil cases? *Daubert* requires trial judges in both civil and criminal proceedings to determine “whether the reasoning or methodology underlying the testimony is scientifically valid.”¹⁹ As the relevant research reveals, however, judges are far more willing to fulfill their gatekeeping roles in civil cases than criminal ones.²⁰ Challenges to forensic **evidence** pretrial, including *Daubert* hearings, are rare in the criminal context.²¹ As the NAS Report makes clear, “the vast majority of the *reported* opinions in criminal cases indicate that trial judges rarely exclude or restrict expert testimony offered by prosecutors.”²² The evidentiary standards that apply to expert forensic **evidence** should be identical in civil and criminal proceedings according to the Federal Rules of Evidence and relevant precedent, yet courts rigorously engage in gatekeeping of such **evidence** in civil proceedings while giving broad leeway to prosecutors in criminal proceedings.²³ Therefore, the courts' failure to exclude faulty forensics in criminal cases cannot be explained away simply by pointing to judges' lack of scientific prowess.

*7 Nor can the courts' repeated failure to exclude **unreliable** criminal expert **evidence** be excused by assertions that the type of scientific **evidence** proffered in civil cases is either substantially materially different or easier for judges to evaluate than that propounded in criminal cases. Virtually every imaginable criminal case has a civil analogue, which requires production of the same or similar **evidence** to secure a verdict (albeit under the relaxed preponderance of the **evidence** or clear and convincing **evidence** standards of review).²⁴ Moreover, we argue that, to the extent that there is any material difference in the type of scientific **evidence** propounded between the two types of proceedings, it is civil cases, including products liability and mass toxic tort cases, and not criminal cases that typically present more difficult reliability, validity, and causation questions for courts.²⁵

We further contend that the frequent admission of flawed forensics in criminal cases cannot be blamed solely on prosecution bias or pro-conviction motives. Even a cursory comparison of the criminal and civil pretrial discovery and disclosure rules demonstrate that a systems-level problem is a contributing culprit. While civil defendants have

successfully implored courts to set the bar very high for the admission of scientific **evidence**, such as epidemiological and toxicological causation **evidence**. prosecutors have encouraged courts to readily admit forensic **evidence** that does not withstand scientific scrutiny.

II. PRETRIAL RULES FAVOR THE REPEAT LITIGANT

Certain litigants in both the civil and criminal systems are “repeat players.” Whereas the repeat players in civil litigation are defendant corporate and government entities, the repeat players in the criminal justice *8 system are prosecutors.²⁶ Repeat players influence pretrial adjudication by “advocating for interpretations of rules and decisions that favor long-term litigation objectives.”²⁷ Individual civil plaintiffs and criminal defendants (“one-shotters or OSs”), on the other hand, are incentivized to “seek out ... short-term gain[s] that may on balance harm future civil plaintiffs and criminal defendants,” rather than pursue any long game.²⁸ As Professor Rothstein explains:

The large-volume litigant is able to achieve the most favorable forum; emphasize different issues in different courts; take advantage of differences in procedure among courts at the state and federal levels; drop or compromise unpromising cases without fear of heavy financial loss; stall some cases and push others; and create rule conflicts in lower courts to encourage assumption of jurisdiction in higher courts.²⁹

The key takeaway here is that although repeat litigants are not successful on every position that they advance in court, the sheer volume of litigation that they control allows them to make incremental changes in the law that, over time, amount to considerable long-term advantages.

In that connection, repeat-player civil defendant corporations have made it a priority to enhance judicial scrutiny of scientific forensic **evidence**. In *Daubert* itself, for example, Merrell Dow Pharmaceuticals fought hard to ensure that the jury was precluded from hearing expert epidemiological **evidence** linking its anti-nausea drug, Bendectin, to the young plaintiffs' limb-reduction birth defects.³⁰ In *Joiner*, the General Electric Company similarly battled to exclude plaintiff's expert **evidence** linking his lung cancer to exposure to polychlorinated biphenyls (PCBs) while employed as a company electrician.³¹ Notably, and much like the overwhelming majority of important post-*Daubert* federal appellate decisions, the *Daubert* trilogy³² is comprised exclusively of civil cases involving repeat-player corporate defendants.³³

*9 Neither Federal Rule of Evidence 702 nor *Daubert* distinguish in any manner between civil and criminal cases regarding the admissibility standards that pertain to expert **evidence**.³⁴ Indeed, “**evidence** law to a significant extent was itself a product of treating criminal and civil cases alike [It] has remained unified because the rebuttable presumption has remained that rules of **evidence** should apply ‘across the board.’”³⁵ Nonetheless, judges have “assessed the ‘reliability’ of expert testimony in civil cases much more rigorously than in criminal cases.”³⁶ Since *Daubert*, traditional forms of criminal forensic **evidence**, such as bite marks, handwriting, hair samples, and fingerprints, have been admitted routinely, bypassing the rigorous methodology scrutiny that applies to, for example, epidemiological and toxicological causation **evidence** in civil products liability and toxic tort cases. As one commentator concluded, “[j]udicial scrutiny in civil litigation and judicial passivity in criminal litigation is aligned with the repeat-player dynamic unique to each forum.”³⁷ Corporate defendants in civil cases routinely challenge the faulty forensic **evidence** used against them, pushing judges to be more skeptical in civil proceedings. By contrast, prosecutors consistently introduce the same **evidence** in criminal cases, encouraging judges in criminal proceedings to rely on precedent. Over time, this has created

DISCOVERING FORENSIC FRAUD, 112 Nw. U. L. Rev. Online 1

a discrepancy in how trial judges rule on scientific **evidence** in civil versus criminal settings that cannot be explained by a difference in substantive law or the applicable rules of **evidence**.

III. PRETRIAL DISCOVERY IN CRIMINAL AND CIVIL PROCEEDINGS

The Federal Rules of Civil Procedure mandate that all parties freely exchange information, including the disclosure of any expert **evidence** throughout the pretrial proceedings.³⁸ By contrast, prosecutors are required to provide criminal defendants very limited pretrial discovery. The Federal Rules of Criminal Procedure, for example, do not entitle a criminal defendant to review either his grand jury transcript or any of the **evidence** the government presented to the grand jury.³⁹ The government does not have to provide the defendant any statements made by government ***10** attorneys or any of its witnesses, including law enforcement agents.⁴⁰ Neither the government nor the accused is subject to any automatic disclosure requirements except the prosecutor's *Brady v. Maryland*⁴¹ duty to produce exculpatory **evidence**.⁴² Moreover, discovery depositions are nonexistent in the criminal justice system. Indeed, criminal depositions are permitted exclusively to preserve the testimony of a party's own witness who may be unavailable for trial.⁴³

By contrast, open and mandatory disclosure of proffered scientific expert **evidence** pretrial in the civil system has had a significant impact on the quality of forensic **evidence**, generally, and causation **evidence**, specifically, that a civil plaintiff must proffer to survive a *Daubert* challenge. As Professor Joseph Sanders explains, “[i]n no area [of the law] has the *Daubert* revolution had a greater effect than in [civil] toxic torts. The number of cases in which expert causation testimony has been excluded must by now run into the thousands.”⁴⁴ In marked contrast to the criticism surrounding courts' routine admission of questionable criminal forensic **evidence**, “[m]any commentators have reacted negatively to this trend [of excluding general causation **evidence** in civil cases], arguing that the bar has been set *too high*.”⁴⁵ Regardless of whether one agrees that the admissibility standards applicable to general causation **evidence** in civil cases strike the right balance, it is widely acknowledged that the predominant exclusionary decisions have forced toxic tort and products ***11** liability plaintiffs to proffer high quality scientific **evidence** to survive pretrial *Daubert* challenges.⁴⁶

IV. LACK OF PRETRIAL DISCOVERY IN CRIMINAL PROCEEDINGS CONTRIBUTES TO THE CONTINUED ADMISSION OF FAULTY FORENSICS

The lack of discovery of scientific **evidence** pretrial in the criminal justice system both affects individual cases and contributes to the culture of admission particular to certain forensic disciplines. The PCAST Report highlighted the need for increased rigor in assessing the scientific validity of **evidence** from a variety of forensic disciplines, many of which employ feature-comparison methodologies, including hair, latent fingerprint, firearm, **DNA** complex-mixture sample, footwear, and bite mark analysis.⁴⁷ As the Report frankly explains, “reviews by competent bodies of the scientific underpinnings of forensic disciplines and the use in courtrooms of **evidence** based on those disciplines have revealed a dismaying frequency of instances of use of forensic **evidence** that do not pass an objective test of scientific validity.”⁴⁸

Bite mark **evidence**, otherwise known as forensic odontology, has been the subject of significant scrutiny. Forensic odontology entails examining marks left on skin or an object to determine if they are human bite marks and then comparing those human bite marks to a suspect's dental impressions.⁴⁹ Not only has the discipline proven incapable of reliably individuating an alleged bite mark--that is, establishing that a bite mark belongs to a specific individual--it cannot even reliably identify skin marks as either human or animal bite marks.⁵⁰ As recently as the spring of 2015, the American Board of Forensic Odontology (ABFO) was unable to find consensus among thirty-nine ABFO-certified bite mark

DISCOVERING FORENSIC FRAUD, 112 Nw. U. L. Rev. Online 1

experts on whether a patterned injury was a human bite mark or if it had identifying features for individualization.⁵¹ In the same year, the Assistant Director of *12 the White House Office of Science and Technology Policy singled out bite mark **evidence** as an example of an **unreliable** forensic discipline and called for its “eradication.”⁵²

Shockingly, courts continue to admit bite mark **evidence** in criminal trials and do so virtually exclusively on the bases of precedent. Demonstrating the powerful influence of the repeat litigant prosecutor, courts continue to admit prosecutor's proffers of **unreliable** bite mark **evidence** in criminal cases, notwithstanding the fact that “bite mark **evidence** has led to more than two dozen wrongful arrests or convictions.”⁵³ Indeed, admitting courts mistakenly rely on prosecutorial arguments that bite marks have been accepted as a valid scientific theory by a sister court instead of conducting an independent *Daubert* analysis.⁵⁴ The treatise on Modern Scientific **Evidence** itself states that “rather than the field [of forensic odontology] convincing the courts of the sufficiency of its knowledge and skills, admission by the courts seems to have convinced the forensic odontology community that, despite their doubts, they were indeed able to perform bite mark identifications after all.”⁵⁵

Worse yet, courts have justified their admission of bite mark **evidence** by relying on certain bite mark cases that resulted in wrongful convictions.⁵⁶ In *State v. Armstrong*, the West Virginia Supreme Court of Appeals took judicial notice of the “general acceptance” of bite mark **evidence**, provoking a cascade of similar court rulings.⁵⁷ The *Armstrong* Court, however, had relied on the Wisconsin case of Robert Lee Stinson, who was ultimately exonerated of his crime in 2009 through **DNA evidence**,⁵⁸ in reaching that conclusion.

Notwithstanding this admonition, not a single federal or state criminal court has upheld a challenge to exclude bite mark **evidence** to date.⁵⁹ Instead, the only serious evaluation of bite mark **evidence** by courts has *13 occurred in civil post-conviction habeas corpus cases and 42 U.S.C. § 1983 lawsuits for wrongful conviction and presentation of false **evidence** at trial.⁶⁰ The lack of analysis by criminal trial courts in this context is particularly disheartening given that one of the rationales for replacing the *Frye v. United States*⁶¹ general acceptance rule with the *Daubert* analysis was the notion that certain types of **evidence** offered as “knowledge” frequently creep into general acceptance without any careful examination of its scientific reliability and validity and “[t]his is especially likely to be true of knowledge that has been widely accepted for a considerable time.”⁶²

V. SOLUTION: PRETRIAL DISCOVERY AND DISCLOSURE IN CRIMINAL PROCEEDINGS

As explained above, “[i]n civil cases and especially tort cases, judges ... enforce *Daubert* aggressively and often insightfully, showing considerable acumen about research methodology.”⁶³ Indeed, “[i]n federal courts, where the decision is legally binding, *Daubert* has become a potent weapon of tort reform by causing judges to scrutinize [civil] scientific *14 **evidence** more closely.”⁶⁴ As a result, the authors endorse the adoption of federal civil pretrial discovery and disclosure procedure in criminal cases. We are not alone. In the wake of the public revelations of wrongful convictions in their respective states, Texas, North Carolina, and West Virginia have reformed their criminal discovery standards to provide pre-plea disclosure of **evidence** to the defendant.⁶⁵

Alan Gell was freed from North Carolina's death row because the prosecution suppressed, throughout his trial proceedings, significant exculpatory and impeachment **evidence**, including the statements of seventeen separate witnesses, each of whom saw the victim alive after Mr. Gell was incarcerated.⁶⁶ In response, North Carolina adopted open criminal

DISCOVERING FORENSIC FRAUD, 112 Nw. U. L. Rev. Online 1

discovery in 2004.⁶⁷ In 2011, the state's legislature enacted the Forensic Sciences Act, which automatically requires law enforcement officers and crime labs--investigative agencies under the wing of the prosecution--to disclose **evidence** to the defense.⁶⁸ The Act also criminalized the failure of law enforcement to disclose scientific **evidence**, including analyst working papers such as bench notes and preliminary tests, to prosecutors.⁶⁹

Emphasizing investigative agencies' obligation to disclose their own **evidence** to the prosecution is particularly important. In *Kyles v. Whitley*, the United States Supreme Court expanded prosecutorial *Brady* obligations by holding that prosecutors have an affirmative duty to disclose favorable **evidence** to the defense, including **evidence** in the hands of the police unknown to the prosecutor.⁷⁰ After the Supreme Court reversed Mr. Kyles's conviction, the prosecution retried him three times, resulting in three hung juries. More pertinently, the prosecution provided previously undisclosed and material police **evidence** to the defense at each of these retrials.⁷¹

In West Virginia, Joseph Buffey pled guilty to rape and burglary while prosecutors were in possession of exculpatory **DNA evidence**.⁷² Mr. Buffey *15 spent the next thirteen years attempting to retract his guilty plea, which local prosecutors uniformly resisted. The West Virginia Supreme Court of Appeals ultimately allowed Mr. Buffey to rescind his guilty plea, and ruled that all state prosecutors must disclose exculpatory **evidence** to criminal defendants pre-plea.⁷³ Accordingly, West Virginia--the same state that judicially noticed bite mark **evidence**--requires the prosecution to disclose *Brady evidence* to the defense during plea negotiations. Notably, in a concurrence in the *Buffey* decision, Justice Allen Hays Loughry stated, “[t]here is simply no room in our judicial system for unethical evidentiary gamesmanship.”⁷⁴

In Texas, Michael Morton was wrongfully convicted of his wife's murder after his prosecutor--who later became a judge--hid exculpatory **evidence**.⁷⁵ The Texas legislature responded by passing the Michael Morton Act, which requires full open-file discovery of favorable **evidence** “as soon as practicable” after the prosecution receives a request.⁷⁶

These states range in their definitions of what constitutes “open-file discovery” from exculpatory **evidence** only in West Virginia to all **evidence** in the prosecutor's file in North Carolina. In all six states with open-discovery provisions, the prosecution is required to disclose--at a minimum--**evidence** favorable to the defense pretrial.⁷⁷ Generally, open-file discovery means the defendant is entitled to the complete file of the prosecution, law enforcement, and any other agencies working for the prosecution. The term “file” broadly includes “witness statements, investigating officers' notes, results of [forensic] tests and examinations,” bench notes and working papers from forensic lab analysts, forensic expert reports, and any other forensic **evidence** collected during the investigation.⁷⁸ Consistent with the position taken by the American Bar Association, open-file states generally require prosecutors to disclose all **evidence** related to a case pre-plea.⁷⁹

*16 The purpose of open-file discovery is to increase the reliability and accuracy of criminal proceedings. As eloquently stated by Professor Robert Mosteller, “[open files] do not rely on the ethical judgment of a prosecutor involved in a fiercely competitive adversary trial process to determine what is exculpatory. Instead, they impose a blanket rule of general disclosure.”⁸⁰ The Honorable Alex Kozinski, Judge on the United States Court of Appeals for the Ninth Circuit, and Senior Advisor to PCAST, likewise suggests open-file discovery as a reform for prosecutorial misconduct.⁸¹ As Professor Jennifer Laurin has made clear, “[e]xpanding and accelerating defense access to information adduced in the state's investigation is one of the most promising mechanisms to remedy reliability-diminishing features of pretrial activities.”⁸²

DISCOVERING FORENSIC FRAUD, 112 Nw. U. L. Rev. Online 1

And yet, even if the Supreme Court had ruled in *Turner* that *Brady* was broad enough to demand prosecutorial disclosure of the alternative perpetrator witness statements to the defense, which it did not, *Brady* would remain an insufficient safeguard and continue to fall far short of the civil discovery rules.⁸³ Despite *Brady*'s narrow scope, the Department of Justice has strongly resisted the incorporation of *Brady* and its progeny into [Federal Rule of Criminal Procedure 16](#). Needless to say, the Department has vehemently opposed the adoption of a parity-based open discovery and disclosure system comparable to those mandated by the Federal Rules of Civil Procedure. Thus, pretrial discovery and disclosure available to federal defendants remain extremely limited, the ABA's proposed reforms and the recent evolution of state rules toward open-file criminal discovery notwithstanding.⁸⁴

***17 CONCLUSION**

This Essay responds to a critical situation in our modern criminal justice system: the ongoing and affirmative use of flawed forensic **evidence** by prosecutors. We have taken this opportunity to identify an underlying systemic issue of discovery by comparing the lax admission standards of false scientific **evidence** in criminal cases with the rigorous vetting of even valid and reliable scientific **evidence** in the civil context. In both criminal and civil cases, the same **evidence** is reviewed by the same judges applying the same standard of admission of scientific **evidence**: *Daubert*. The difference, and one that undermines the accuracy not only of the **evidence** presented but also of criminal convictions, is the pretrial discovery and disclosure rules binding the courtroom players. We propose that the criminal justice system adopt the party-parity civil pretrial discovery and disclosure rules. Such leveling of the playing field may return integrity to prosecutors' offices and restore trust in our criminal adjudications.

Footnotes

- ^{a1} Jennifer D. Oliva, Associate Professor of Law and Public Health, West Virginia University; Valena E. Beety, Associate Professor of Law, West Virginia University College of Law. We thankfully acknowledge helpful comments from Brandon Garrett, Adam Shniderman and Edward Cheng. We also appreciated the opportunity to present this piece at the 2017 International Forensic Science Error Management Symposium at the National Institute of Standards and Technology (NIST). Our work on this project was supported by a generous Hodges Research Grant from West Virginia University College of Law.
- ^{d1} AN ADDENDUM TO THE PCAST REPORT ON FORENSIC SCIENCE IN CRIMINAL COURTS 4, 9 (2017), https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensics_addendum_finalv2.pdf [<https://perma.cc/HQW9-FHRU>].
- ¹ 137 S. Ct. 1885 (2017).
- ² *Brady v. Maryland*, 373 U.S. 83 (1963).
- ³ PRESIDENT'S COUNCIL OF ADVISORS ON SCI. & TECH., EXEC. OFFICE OF THE PRESIDENT, FORENSIC SCIENCE IN CRIMINAL COURTS: ENSURING SCIENTIFIC VALIDITY OF FEATURE-COMPARISON METHODS (2016), https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf [<https://perma.cc/XFS5-96E2>] [hereinafter PCAST Report].
- ⁴ See, e.g., ALLIANCE FOR JUSTICE, BROADENING THE BENCH: PROFESSIONAL DIVERSITY AND JUDICIAL NOMINATIONS 5-6 (2016), <http://www.afj.org/wp-content/uploads/2014/11/Professional-Diversity-Report.pdf> [<https://perma.cc/TF59-GEAB>] (explaining that, of President Obama's federal judicial appointees, “[p]rosecutors outnumber public defenders (state or federal) by three to one; [o]nly five out of 64 circuit nominees have worked as a public defender (state or federal), compared to 24 who have worked as prosecutors; [and a]pproximately 86% have been either corporate attorneys or prosecutors (or both)”); Editorial, *The Homogeneous Federal Bench*, N.Y. TIMES (Feb. 6, 2014), <https://>

mobile.nytimes.com/2014/02/07/opinion/the-homogeneous-federal-bench.html [https://perma.cc/PLV6-G3AK] (“[U]nder the Obama administration, federal judges continue to be drawn overwhelmingly from the ranks of prosecutors and corporate lawyers. This deprives the courts of crucial perspectives and reduces public trust in the justice system.”); *see also* Dara Lind, *There Hasn't Been a Criminal Defense Lawyer on the Supreme Court in 25 Years. That's a Problem.*, VOX (Mar. 22, 2017), <https://www.vox.com/2016/3/28/11306422/supreme-court-prosecutors-career> [https://perma.cc/FNY6-NGPS] (noting that while there are no former criminal defense attorneys on the Supreme Court, there are three ex-prosecutors); Nicki Gorny, *Pipeline to the Bench: New Judges Often Former Prosecutors*, OCALA STARBANNER (Nov. 14, 2015), <http://www.ocala.com/news/20151114/pipeline-to-the-bench-new-judges-often-former-prosecutors> [https://perma.cc/BB3E-ZUDW] (stating “[i]f you commit a crime in Marion County[, Florida,] next year, there's a 3 in 4 chance that you'll face a former prosecutor on the bench”).

5 Ion Meyn, *The Unbearable Lightness of Criminal Procedure*, 42 AM. J. CRIM. L. 39, 47 (2014); *see also* Brandon L. Garrett, *Aggregation in Criminal Law*, 95 CAL. L. REV. 383, 400 (2007) (“In our current criminal system, repeat players generate the means to achieve vast economies of scale resulting in fewer criminal trials and therefore fewer opportunities to vindicate criminal procedural rights at trial.”).

6 *See, e.g.*, Meyn, *supra* note 5, at 46-47; Jennifer E. Laurin, *Quasi-Inquisitorialism: Accounting for Deference in Pretrial Criminal Procedure*, 90 NOTRE DAME L. REV. 783, 785 (2014) (explaining that “the structure of American criminal procedure doctrine ... relies almost entirely on trial-based procedures to guarantee accuracy and approaches the pretrial realm with a comparatively light regulatory touch”).

7 Marc Galanter, *The Vanishing Trial: An Examination of Trials and Related Matters in Federal and State Courts*, 1 J. EMPIRICAL & LEGAL STUD. 459, 459 (2004); *see also* Joseph F. Anderson, Jr., *Where Have You Gone, Spot Mozingo? A Trial Judge's Lament over the Demise of the Civil Jury Trial*, 4 FED. CTS. L. REV. 99, 101 (2010) (discussing “the vanishing jury trial”); William G. Young, *Vanishing Trials, Vanishing Juries, Vanishing Constitution*, 40 SUFFOLK U. L. REV. 67, 73 (2006) (acknowledging that the “civil jury trial has all but disappeared”) (quoting Kevin M. Clermont & Theodore Eisenberg, *Litigation Realities*, 88 CORNELL L. REV. 119, 142-43 (2002); Lawrence M. Friedman, *The Day Before Trials Vanished*, 1 J. EMPIRICAL & LEGAL STUD. 689, 691 (2004) (“By the end of the 19th century, it was already the case that the vast majority of convictions in felony cases came about as a result of a guilty plea.”).

8 As the Supreme Court has aptly recognized, due to the civil discovery rules, “civil trials in the federal courts no longer need be carried on in the dark. The way is now clear ... for the parties to obtain the fullest possible knowledge of the issues and facts before trial.” *Hickman v. Taylor*, 329 U.S. 495, 501 (1947); *see also* *United States v. Proctor & Gamble Co.*, 356 U.S. 677, 682 (1958) (“Modern instruments of discovery ... [and] pretrial procedures make a trial less a game of blindman's [sic] bluff and more a fair contest with the basic issues and facts disclosed to the fullest practicable extent.”).

9 Georgia A. Staton & Renee J. Scatena, *Parallel Proceedings--A Discovery Minefield*, 34 ARIZ. ATTY 17, 18 (1998) (noting that “[t]he absence of mandatory disclosure and the limited permissive disclosure provisions increase the investigative burden on the criminal defendant. The prosecution, with its abundant resources and access to federal agents, holds the advantage.”); Meyn, *supra* note 5, at 41 (explaining that “[t]he absurd result is that the class of litigants traditionally warranted robust protection receives the least protection”).

10 *See generally* Jennifer L. Groscup et al., *The Effects of Daubert on the Admissibility of Expert Testimony in State and Federal Criminal Cases*, 8 PSYCHOL. PUB. POL'Y & L. 339 (2002); *see also* *United States v. Sherwood*, 98 F.3d 402, 408 (9th Cir. 1996) (admitting fingerprint comparison **evidence** without conducting a *Daubert* hearing). In *United States v. Havvard*, the court described *Sherwood* as an opinion “asserting that *the reliability of fingerprint comparisons cannot be questioned.*” 260 F.3d 597, 600 (7th Cir. 2001) (emphasis added).

11 GA. CODE ANN. § 24-7-707 (2016) (“[T]he opinions of experts on any question of science, skill, trade, or like questions shall always be admissible”).

DISCOVERING FORENSIC FRAUD, 112 Nw. U. L. Rev. Online 1

- 12 The Georgia legislature has adopted standards applicable to its civil expert witnesses that are nearly identical to those provided by the Federal Rules of **Evidence** and Civil Procedure. See *GA. CODE ANN. § 24-7-702* (2016).
- 13 D. Michael Risinger, *Navigating Expert Reliability: Are Criminal Standards of Certainty Being Left on the Dock?*, 64 *ALB. L. REV.* 99, 100 (2000); see also David A. Sklansky & Stephen C. Yeazell, *Comparative Law Without Leaving Home: What Civil Procedure Can Teach Criminal Procedure, and Vice Versa*, 94 *GEO. L. J.* 683, 714-15 (2005) (explaining that “[c]ivil litigators who venture into criminal cases tend to be stunned and often outraged by their inability to depose government witnesses or even to file interrogatories or requests for admissions”).
- 14 PCAST Report, *supra* note 3.
- 15 NAT'L RES. COUNCIL, NAT'L ACAD. SCI., *STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD* (2009), <https://www.ncjrs.gov/pdffiles1/nij/grants/228091.pdf> [<https://perma.cc/Z9VR-ADYV>] [hereinafter *NAS REPORT*]. The disciplines analyzed by the *NAS REPORT* were biological **evidence** (**DNA** analysis), controlled substances analysis, fingerprints (friction ridge analysis), pattern/impression **evidence**, tool mark and firearm identification, hair analysis, fiber **evidence** analysis, questioned document examination, paint and coatings analysis, explosives and fire analysis, forensic odontology (bite marks), bloodstain and pattern analysis, and digital and multimedia analysis. See also NAT'L ASS'N OF CRIM. DEF. LAWS., *PRINCIPLES AND RECOMMENDATIONS TO STRENGTHEN FORENSIC EVIDENCE AND ITS PRESENTATION THE COURTROOM 8* (2010), <https://www.nacdl.org/WorkArea/DownloadAsset.aspx?id=21802> [<https://perma.cc/6HBM-XDVR>] (recommending that “[t]he results of any forensic theory or technique whose validity, limitations, and measures of uncertainty have not been established should not be admitted into **evidence** to prove the guilt of an accused person”).
- 16 PCAST Report, *supra* note 3, at 40-43.
- 17 *White House Advisory Council Report is Critical of Forensics Used in Criminal Trials: U.S. Attorney General Says Justice Department Won't Adopt Recommendations*, *WALL ST. J.* (last updated Sept. 20, 2016, 4:25 PM), <http://www.wsj.com/articles/white-house-advisory-council-releases-report-critical-of-forensics-used-in-criminal-trials-1474394743> [<https://perma.cc/BM3W-M79C>].
- 18 Adam B. Shniderman, *Prosecutors Respond to Calls for Forensic Science Reform: More Sharks in Dirty Water*, 126 *YALE L.J.F.* 348, 352-57 (2017).
- 19 *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 592-93 (1993).
- 20 Risinger, *supra* note 13, at 99 (explaining that “as to proffers of asserted expert testimony, civil defendants win their *Daubert* reliability challenges to plaintiffs’ proffers most of the time, and that criminal defendants virtually always lose their reliability challenges to government proffers”).
- 21 Peter J. Neufeld, *The (Near) Irrelevance of Daubert to Criminal Justice and Some Suggestions for Reform*, 95 *AM. J. PUB. HEALTH* S107, S107 (2005); Erica Beecher-Monas, *Blinded by Science: How Judges Avoid the Science in Scientific **Evidence***, 71 *TEMP. L. REV.* 55, 56-57 (1998).
- 22 *NAS REPORT*, *supra* note 15, at 11.
- 23 Shniderman, *supra* note 18, at 354.
- 24 For example, “[a]ll states provide for a [civil] cause of action for wrongful death by a Wrongful Death Statute.” Jay W. Elston, *State Wrongful Death Acts and Maritime Torts*, 39 *TEX. L. REV.* 643, 645 (1961); see also, e.g., Sklansky & Yeazell, *supra* note 13, at 687 (explaining that “[a]s recently as the nineteenth century--indeed, well into the twentieth century--civil and criminal proceedings were, in essence, alternative ways for aggrieved victims of wrongs to enlist the adjudicative machinery of the state in seeking redress”).

DISCOVERING FORENSIC FRAUD, 112 Nw. U. L. Rev. Online 1

- 25 Dáirdre Dwyer, (*Why*) *are Civil and Criminal Expert Evidence Different?*, 43 TULSA L. REV. 381, 387-88 (2007) (explaining the uniqueness of epidemiological evidence of causation in toxic tort to civil proceedings and positing that such evidence “has a high scientific content, and the demonstration of causation is indirect in that it rests on arguments about whether the claimant was statistically more likely to suffer harm as a result of exposure to the allegedly toxic substance. The scientific evidence has not been collected to address directly the question of whether a specific individual has suffered harm”); *see also* Risinger, *supra* note 13, at 102 (explaining that “[i]t is unlikely to be pure coincidence that the Supreme Court chose a civil case, *Daubert v. Merrell Dow Pharmaceuticals Inc.*, to review the appropriate criteria of dependability, or that its two subsequent forays into these waters have also been in civil cases”) (footnote omitted).
- 26 Marc Galanter, *Why the “Haves” Come Out Ahead: Speculations on the Limits of Legal Change*, 9 LAW & SOC’Y REV. 95, 97 (1974) (explaining that “[t]he spouse in a divorce case, the auto-injury claimant, the criminal accused are OSs [one-shotters]; the insurance company, the prosecutor, the finance company are RPs [repeat players]”).
- 27 Meyn, *supra* note 5, at 47.
- 28 *Id.*
- 29 Lawrence E. Rothstein, *The Myth of Sisyphus: Legal Services Efforts on Behalf of the Poor*, 7 U. MICH. J. L. REFORM 493, 501 (1974).
- 30 *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 583-85 (1993).
- 31 *Gen. Elec. Co. v. Joiner*, 522 U.S. 136, 136 (1997).
- 32 The *Daubert* trilogy includes *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993), *General Electric Co. v. Joiner*, 522 U.S. 136 (1997), and *Kumho Tire Co. v. Carmichael*, 526 U.S. 137 (1999).
- 33 Neufield, *supra* note 21, at S109 (explaining that “it is not a coincidence that ... almost all of the post-*Daubert* federal appellate decisions that further defined the standard have been civil rather than criminal”).
- 34 Sklansky & Yeazell, *supra* note 13, at 730-31.
- 35 *Id.* at 728, 730.
- 36 *Id.* at 731.
- 37 Meyn, *supra* note 5, at 48 (internal citations omitted).
- 38 *See, e.g.*, FED. R. CIV. P. 26. Expert evidence must be disclosed pretrial pursuant to Federal Rules of Evidence 702, 703 and 706.
- 39 FED. R. CRIM. P. 16(a)(3).
- 40 FED. R. CRIM. P. 16(a)(2). Under the Jencks Act, 18 U.S.C. § 3500, the government’s witness statements are only discoverable by the defense after the witness has testified on direct examination and after the defense has properly requested the statements. *See Jencks v. United States*, 353 U.S. 657, 67071 (1957).
- 41 373 U.S. 83, 88 (1963).
- 42 Notably, the prosecution is not required to disclose *Brady* material pre-plea so long as other due process protections are in place. *United States v. Ruiz*, 536 U.S. 622, 631 (2002) (explaining that where government was required to give defendant information regarding factual innocence before plea no other *Brady* disclosure was required).
- 43 FED. R. CRIM. P. 15(a)(1). As Professor Meyn recently explained, “[t]he resistance to granting a criminal defendant the power to investigate has deep roots.” Ion Meyn, *Discovery and Darkness: The Information Deficit in Criminal Disputes*, 79 BROOK.

L. REV. 1091, 1120 (2014). The historical arguments against extending formal pretrial discovery to criminal defendants include concerns that such a levelling of the pretrial investigatory playing field would give criminal defendants an unfair advantage, enable them to threaten and intimidate witnesses, and lead to the misuse formal powers. *Id.* at 1127-33. Additional anti-reform arguments include allegations that the trial is proper testing of a criminal case, criminal defendants already have enough rights, and extension of formal discovery to criminal defendants would be too costly. *Id.* at 1133-38.

44 Joseph Sanders, *Proof of Individual Causation in Toxic Tort and Forensic Cases*, 75 BROOK. L. REV. 1367, 1374 (2010).

45 *Id.* (emphasis added); see also Carol Krafka et al., *Judge and Attorney Experiences, Practices, and Concerns Regarding Expert Testimony in Federal Civil Trials*, 8 PSYCHOL. PUB. POLY & L. 309, 322 (2002) (noting that judges presiding over civil cases “reported that they were more likely to scrutinize expert testimony before trial and were less likely to admit it” post-*Daubert*).

46 LLOYD DIXON & BRIAN GILL, CHANGES IN THE STANDARDS FOR ADMITTING EXPERT EVIDENCE IN FEDERAL CIVIL CASES SINCE THE DAUBERT DECISION xv (2001), http://www.rand.org/content/dam/rand/pubs/monograph_reports/2005/MR1439.pdf [<https://perma.cc/AG72-CA9E>] (detailing that “[federal] judges scrutinized reliability more carefully and applied stricter standards in deciding whether to admit expert evidence” post-*Daubert*).

47 PCAST REPORT, *supra* note 3.

48 *Id.* at 22.

49 *Id.* at 8.

50 Mary A. Bush et al., *Inquiry into the Scientific Basis for Bitemark Profiling and Arbitrary Distortion Compensation*, 55 J. FORENSIC SCI. 976-83 (2010).

51 Radley Balko, *A Bite Mark Matching Advocacy Group Just Conducted a Study that Discredits Bite Mark Evidence*, WASH. POST (Apr. 8, 2015), <http://www.washingtonpost.com/news/the-watch/wp/2015/04/08/a-bite-mark-matching-advocacy-group-just-conducted-a-study-that-discredits-bite-mark-evidence/> [<https://perma.cc/E7PK-6MP6>] [hereinafter Balko, *Bite Mark Matching Advocacy Group*].

52 Radley Balko, *A High-Ranking Obama Official Just Called for the “Eradication” of Bite Mark Evidence*, WASH. POST (Jul. 22, 2015), <https://www.washingtonpost.com/news/the-watch/wp/2015/07/22/a-high-ranking-obama-official-just-called-for-the-eradication-of-bite-mark-evidence/> [<https://perma.cc/D6YZ-67XF>] [hereinafter Balko, *High-Ranking Obama Official*].

53 Radley Balko, *Incredibly, Prosecutors are Still Defending Bite Mark Evidence*, WASH. POST (Jan. 30, 2017), <https://www.washingtonpost.com/news/the-watch/wp/2017/01/30/incredibly-prosecutors-are-still-defending-bite-mark-evidence/> [<https://perma.cc/QCV9-JA5E>] [hereinafter Balko, *Prosecutors*].

54 Michael J. Saks et al., *Forensic Bitemark Identification: Weak Foundations, Exaggerated Claims*, 3 J.L. & BIOSCI. 538, 546 (2016) (explaining that, in *Burke v. Town of Walpole*, 2004 WL 502617 (D. Mass. 2004), *aff’d in part, vacated in part*, 405 F.3d 66 (1st Cir. 2005), “the federal magistrate judge appeared never to doubt the validity of bite mark expertise though the best the court could do to support its faith was to cite cases that cite cases that express the same credulousness”).

55 4 DAVID L. FAIGMAN ET AL., MOD. SCI. EVIDENCE § 35:4. *The Judicial Response to Expert Testimony on Bitemark Identification*, (2016-2017).

56 M. Chris Fabricant & Tucker Carrington, *The Shifted Paradigm: Forensic Science’s Overdue Evolution from Magic to Law*, 4 VA. J. CRIM. L. 1, 42 & n.173 (2016).

57 *State v. Armstrong*, 369 S.E.2d 870, 877 (W. Va. 1988).

58 *State v. Stinson*, 397 N.W.2d 136 (Wis. Ct. App. 1986).

59 Balko, *High-Ranking Obama Official*, *supra* note 52.

DISCOVERING FORENSIC FRAUD, 112 Nw. U. L. Rev. Online 1

- 60 See, e.g., *Keko v. Hingle*, 318 F.3d 639, 644 (5th Cir. 2003) (denying absolute immunity to forensic odontologist in § 1983 civil lawsuit following wrongful conviction); *Ege v. Yukins*, 380 F. Supp. 2d 852, 871 (E.D. Mich. 2005), *aff'd in part, rev'd in part on other grounds*, 485 F.3d 364 (6th Cir. 2007) (ruling “there is no question that the [bite mark] **evidence** in this case was **unreliable** and not worthy of consideration by a jury”); *In re Richards*, 63 Cal. 4th 291, 315 (2016) (court granting civil writ of habeas corpus ruling bite mark expert's criminal trial testimony constituted material false **evidence**); *Stinson v. Milwaukee*, 2013 WL 5447916, at *12-13 (E.D. Wis. 2013) (denying absolute immunity to forensic odontologists in § 1983 civil lawsuit alleging fabrication and suppression of **evidence**) *aff'd in part, rev'd in part*, *Stinson v. Gauger*, 799 F.3d 833 (7th Cir. 2015).
- 61 293 F. 1013 (D.C. Cir. 1923). In *Frye v. United States*, the District of Columbia Court of Appeals affirmed the trial court's exclusion of expert testimony regarding an early version of a systolic blood pressure-based lie detector test. *Id.* at 1014. The *Frye* Court famously held that “while courts will go a long way in admitting expert testimony deduced from a well-recognized scientific principle or discovery, the thing from which the deduction is made must be sufficiently established to have gained *general acceptance in the particular field in which it belongs*.” *Id.* (emphasis added). “The *Frye* Standard was extremely administrable given that the presiding judge did not need to understand the theories supporting the scientific testimony at hand; he only needed to determine whether the scientific community had accepted the supporting theories as valid.” Claire R. Rollor, *Logic, Not **Evidence**, Supports a Change in Expert Testimony Standards: Why Evidentiary Standards Promulgated by the Supreme Court for Scientific Expert Testimony are Inappropriate and Inefficient When Applied in Patent Infringement Suits*, 8 J. BUS. & TECH. L. 313, 326 (2013). For an extensive discussion of *Frye* and its application to the admission of novel expert **evidence**, see Paul C. Giannelli, *The Admissibility of Novel Scientific **Evidence**: Frye v. United States, A Half-Century Later*, 80 COLUM. L. REV. 1197 (1980).
- 62 David L. Faigman et al., *Check Your Crystal Ball at the Courthouse Door, Please: Exploring the Past, Understanding the Present, and Worrying about the Future of Scientific **Evidence***, 15 CARDOZO L. REV. 1799, 1811 & n.37 (1994) (citing Philip H. Abelson, *The Need for Skepticism*, 138 SCI. 75 (1962)).
- 63 Michael J. Saks, *Judging Admissibility*, 35 J. CORP. L. 135, 144 (2009).
- 64 Edward K. Cheng & Albert H. Yoon, *Does Frye or Daubert Matter? A Study of Scientific Admissibility Standards*, 91 VA. L. REV. 471, 472 (2005).
- 65 See N.C. GEN. STAT. § 15A-903 (2017); 2004 N.C. Sess. Laws 515; TEX. CODE CRIM. PROC. ANN. art. 39.14 (West 2015); *Buffey v. Ballard*, 782 S.E.2d 204, 218 (W. Va. 2015).
- 66 Alexandra Gross, *Alan Gell*, THE NAT'L REGISTRY OF EXONERATIONS, <http://www.law.umich.edu/special/exoneration/Pages/casedetail.aspx?caseid=3236> [<https://perma.cc/XM5J-FQTN>].
- 67 See 2004 N.C. Sess. Laws 515; see also Robert P. Mosteller, *Exculpatory **Evidence**, Ethics, and the Road to the Disbarment of Mike Nifong: The Critical Importance of Full Open-File Discovery*, 15 GEO. MASON L. REV. 257, 272-76 (2008) (explaining the relationship between the *Gell* case and the subsequent criminal discovery reforms enacted by the North Carolina legislature).
- 68 N.C. GEN. STAT. § 15A-903 (2017).
- 69 *Id.*
- 70 514 U.S. 419, 437 (1995) (finding the prosecutor has an affirmative duty to disclose material **evidence**, including “a duty to learn of any favorable **evidence** known to the others acting on the government's behalf in the case, including the police”).
- 71 *Thinking about Kyles v. Whitley*, PUBLIC DEFENDER DUDE, <http://publicdefenderdude.blogspot.com/2010/05/thinking-about-kyles-v-whitley.html> [<https://perma.cc/VZW3-934Y>] (last accessed Mar. 3, 2017).
- 72 *Buffey v. Ballard*, 782 S.E.2d 204, 208-09 (W. Va. 2015).

DISCOVERING FORENSIC FRAUD, 112 Nw. U. L. Rev. Online 1

- 73 *Id.* at 216, 221.
- 74 *Id.* at 223 (Loughry, J., concurring).
- 75 *Michael Morton*, THE NAT'L REGISTRY OF EXONERATIONS, <https://www.law.umich.edu/special/exoneration/Pages/casedetail.aspx?caseid=3834> [<https://perma.cc/U75U-RXWH>] (last accessed Sept. 7, 2017).
- 76 2013 Tex. Sess. Law Serv. Ch. 49 (S.B. 1611) (West) (codified as amended at [TEX. CODE CRIM. PROC. ANN. art. 39.14](#) (West 2017)).
- 77 *See* [N.C. GEN. STAT. ANN. § 15A-903](#) (2011); [TEX. CODE CRIM. PROC. ANN. art. 39.14](#) (West 2017); [ARIZ. R. CRIM. P. 15.1](#); [COLO. CRIM. P. 16](#); [N.J. CT. R. 3:13-3](#); [OHIO CRIM. R. 16](#).
- 78 *See, e.g.*, [N.C. GEN. STAT. ANN. § 15A-903](#) (2011).
- 79 ABA COMM. ON ETHICS AND PROF'L RESPONSIBILITY, Formal Op. 454 (2009) (ethics opinion interpreting ABA Model Rule 3.8(d)).
- 80 Mosteller, *supra* note 67, at 260, 310 (“The beauty of full open-file discovery is obvious as a remedy for the difficulty of subjective choice in a competitive adversarial environment.”).
- 81 Hon. Alex Kozinski, *Criminal Law 2.0*, 44 GEO. L.J. ANN. REV. CRIM. PROC. iii, xxvi-xxvii (2015).
- 82 Laurin, *supra* note 6, at 842.
- 83 It is worth emphasizing here that the authors do not endorse the notion that the adoption of *Brady* and its progeny in [Federal Rule of Criminal Procedure 16](#) would be sufficient to curtail the admission of flawed forensic **evidence** in criminal proceedings. As one scholar has aptly summarized:
[C]onstitutional rules that would subject the prosecutor's (and police's) actions to the scrutiny of the defense--in particular the rule of *Brady v. Maryland* and its progeny entitling the defense, as a feature of due process, to favorable information within the control of the state--do little or nothing to cure information asymmetries prior to trial.
Laurin, *supra* note 6, at 794 (internal citations omitted). This is because:
First, the scope of *Brady's* disclosure requirement is formally limited to information both favorable and ‘material’ to the defense--and thus excludes not only information relevant to the prosecution's case more generally, but also ... favorable information incapable by its own force of affecting a juror's judgment. [Second,] “ordinary course due process [does not] require [] the state to make available *potentially* favorable **evidence**--for example, physical **evidence** that, upon forensic analysis, might yield relevant, even exculpatory, conclusions.” [Third, and] “most critically, even information that falls within the ambit of *Brady's* mandate need not, consistent with the Constitution, be disclosed *prior* to trial.
Id. at 794-95.
- 84 At least six states have adopted some version of pretrial open-file discovery. *See* notes 65-79 *supra* and accompanying text. Florida, while not technically providing open file discovery, provides extensive information to the defense as well. [FLA. R. CRIM. P. 3.220](#).

112 NWULRON 1

41 Law & Hum. Behav. 230

Law and Human Behavior

June, 2017

Saul M. Kassin

John Jay College of Criminal Justice

Jeff Kukucka

Towson University

Victoria Z. Lawson

Institute for State and Local Governance of the City University of New York

John DeCarlo

University of New Haven

Copyright © 2016 by American Psychological Association; Saul M. Kassin, Jeff Kukucka, Victoria Z. Lawson, John DeCarlo

POLICE REPORTS OF MOCK SUSPECT INTERROGATIONS: A TEST OF ACCURACY AND PERCEPTION

A 2-phased experiment assessed the accuracy and completeness of police reports on mock interrogations and their effects on people's perceptions. In Phase 1, 16 experienced officers investigated a mock crime scene, interrogated 2 innocent suspects--1 described by the experimenter as more suspicious than the other--and filed an incident report. All 32 sessions were covertly recorded; the recordings were later used to assess the reports. In Phase 2, 96 lay participants were presented with a brief summary of the case and then either read 1 police report, read 1 verbatim interrogation transcript, or listened to an audiotape of a session. Results showed that (a) Police and suspects diverged in their perceptions of the interrogations; (b) Police committed frequent errors of omission in their reports, understating their use of confrontation, maximization, leniency, and false evidence; and (c) Phase 2 participants who read a police report, compared to those who read a verbatim transcript, perceived the process as less pressure-filled and were more likely to misjudge suspects as guilty. These findings are limited by the brevity and low-stakes nature of the task and by the fact that no significant effects were obtained for our suspicion manipulation, suggesting a need for more research. Limitations notwithstanding, this study adds to a growing empirical literature indicating the need for a requirement that all suspect interrogations be electronically recorded. To provide a more objective and accurate account of what transpired, this study also suggests the benefit of producing verbatim transcripts.

Keywords: interrogations, suspects, police reports, accuracy

Supplemental materials: <http://dx.doi.org/10.1037/lhb0000225.supp>

Over the years, research on police interrogations, confessions, and their role in known cases of wrongful conviction has animated calls for reform (for reviews, see Gudjonsson, 2003; Kassin, 1997, 2005, 2012; Kassin & Gudjonsson, 2004; Lassiter & Meissner, 2010; for an official white paper, see Kassin et al., 2010). Many such reform efforts have been aimed at protecting highly vulnerable populations (e.g., juveniles, people with intellectual or mental health impairments) and

at curtailing the use of coercive interrogation practices (e.g., presentations of false evidence, minimization tactics that imply leniency). Perhaps the most significant proposed safeguard is to require the electronic recording of interrogations--the entire process, not just the confession. As stated in the AP-LS white paper: "Without equivocation, our most essential recommendation is to lift the veil of secrecy from the interrogation process in favor of the principle of transparency" (Kassin et al., 2010, p. 25).

There is a perennial debate concerning the recording of suspect interviews and interrogations (for an overview, see Drizin & Reich, 2004). As an historical matter, the practice has drawn strong resistance from many federal, state, and local police professionals (e.g., Boetig, Vinson, & Weidel, 2006)--especially those trained by John Reid & Associates, which, until recently, had steadfastly opposed the recording of interrogations (Inbau, Reid, Buckley, & Jayne, 2001). The bases of opposition have varied. Some have opposed recording on pragmatic and logistical grounds--citing the scope of such a requirement; financial costs; the evidentiary consequences of a failure to comply, for example, due to equipment malfunction; and issues of consent, especially in two-party consent states. Others have expressed concern over how recording might alter the behavior of both police and suspects during interrogation and the subsequent decision-making of judges and juries *231 (for an overview of these arguments, see Sullivan, 2008; Thurlow, 2005).

At the same time, a policy of mandatory recording has received support from various organizations (e.g., American Bar Association, 2004; American Psychological Association, 2014; Buckley & Jayne, 2005; The Justice Project, 2007), as well as from surveys of police investigators across the United States and Canada (Geller, 1993; Kassin et al., 2007). Sullivan (2004) interviewed police from hundreds of departments that recorded custodial interrogations and consistently found that they fully embraced the practice. Respondents cited numerous benefits of recording, namely that it allowed detectives to focus on the suspect rather than on concurrent notetaking; that it allowed them to later review the suspect's answers to questions for any incriminating comments that had initially gone unnoticed; that it lessened the need for detectives to defend their interrogation practices in court; and that it enhanced public trust in law enforcement (also see Sullivan, Vail, & Anderson, 2008). Importantly, the U.S. Department of Justice recently reversed its longstanding opposition to recording by establishing the presumptive requirement that federal law enforcement agencies (including the FBI) videotape the custodial interrogations of felony suspects (Schmidt, 2014).

That police who record interrogations report high levels of satisfaction is a vital data point in efforts to reform majority practice. But what are the actual effects? There are two primary sets of reasons for the proposed recording requirement. The first is the expectation that the practice of recording will increase accountability and discourage the use of coercive interrogation tactics, thus reducing the risk to all suspects. To test this hypothesis, Kassin, Kukucka, Lawson, and DeCarlo (2014) conducted an experiment in a police station, in which 61 experienced investigators interrogated a male suspect who was either guilty or innocent of a mock theft. Before each interrogation, the investigator either was or was not informed that their session would be surreptitiously recorded. These recordings were later coded for the use of various high-pressure tactics designed to elicit a confession. As predicted, camera-informed interrogators were less likely than their uninformed counterparts to use both maximization and minimization tactics; they were also judged by suspects--who were not told of the camera's presence--as trying less hard to obtain a confession. These findings suggest that recording can affect the process of interrogation--namely, by inhibiting the use of coercive tactics.

A second purported benefit of recording interrogations is to provide an accurate factual record of the interrogation behavior of police and suspects. Perhaps the most frequently invoked argument is that it is the most effective way to memorialize the process by which a statement was taken and, hence, increase the fact-finding accuracy of prosecutors (who decide whether to charge a suspect), judges (who rule on whether a confession was voluntary or coerced), and juries (who determine whether a confession is credible and hence whether the confessor is guilty or innocent). In current practice, whereby detectives take contemporaneous or retrospective notes, disputes often arise as to whether *Miranda* rights were administered and waived in a timely manner; whether the suspect was cooperative or evasive; whether police made or implied promises or threats, or lied about evidence; and, importantly, whether the details contained within a confession originated from the suspect. Disputes over this latter issue can prove devastating. In a descriptive analysis of

38 false confessions from the Innocence Project, Garrett (2010) found that 36 were “contaminated,” containing accurate crime details allegedly known only to the perpetrator--details not in the public domain, but known to police, that the innocent confessor could not have produced without exposure to secondhand information. The result: An increased likelihood of conviction, as the presence of details in a confession enhances perceptions of its credibility (Appleby, Hasel, & Kassin, 2013).

In lieu of electronic recordings, the Federal Rules of Evidence (2015) provide that a police witness may use a personally prepared report concerning an interview or interrogation, including what the defendant said, to refresh his or her recollection while testifying (FRE 612). Yet the accuracy of these reports, which is often in dispute, has never been tested in the context of a suspect interrogation. The question we posed in the present research is: What does recording reveal about the memorial accuracy of police reports of interrogations, and with what effect on prospective fact finders?

Basic research on memory for conversation content suggests that this process may be fraught with bias and error. Neisser (1981) highlighted this problem in a case study in which he analyzed John Dean's high-stakes memory of Watergate-related conversations with President Nixon. Dean testified as to his recollections with specificity and confidence. Yet when tapes of oval office conversations were discovered, Neisser discovered that although Dean was generally correct about what happened (e.g., that there was a cover-up), his memory of specifics was often distorted (e.g., overestimating his own role). This case stands in contrast to research showing that active involvement in a conversation tends to enhance memory for content; for example, people recall what they said better than what they read or heard (MacLeod, Gopie, Hourihan, Neary, & Ozubko, 2010).

Apart from the potential for intrusion and omission errors in memory for conversations, research shows that memories of *content* and *context* are stored independently, thereby increasing the risk of source monitoring confusion (Johnson, Hashtroudi, & Lindsay, 1993). Moreover, such errors often reflect the operation of cognitive confirmation biases (Kleider, Pezdek, Goldinger, & Kirk, 2008; Schacter, 2001). In a study that demonstrated the problem in a forensic context, Lamb, Orbach, Sternberg, Hershkowitz, and Horowitz (2000) took a sample of 20 interviews of alleged child sex abuse victims and compared interviewers' contemporaneous notes against audiotapes of these sessions. The interviewers' notes proved inadequate, as they failed to mention 57% of their own utterances and 25% of details that the children provided. Moreover, the notes often contained serious source attribution errors, such as citing the children rather than their own prompting questions as the source of details that were disclosed. Lamb et al. (2000) concluded: “Even when they made contemporaneous verbatim notes, these investigators tended to understate their role in eliciting the information” (p. 705; see also Bruck, Ceci, & Francoeur, 1999).

The consequences of source attribution errors were realized by former D.C. Detective James Trainum (2007) who--in an article entitled “I took a false confession so don't tell me it doesn't happen!”--described a case in which a former suspect who had confessed was later exonerated:

Years later, during a review of the videotapes, we discovered our mistake. We had fallen into a classic trap. We believed so much in our *232 suspect's guilt that we ignored all evidence to the contrary. To demonstrate the strength of our case, we showed the suspect our evidence, and unintentionally fed her details that she was able to parrot back to us at a later time. It was a classic false confession case and without the video we would never have known.

To further complicate matters of recollection, there is reason to believe that suspects likewise cannot be trusted to provide accurate accounts. Unlike most social interactions, police interrogations are exceptionally stressful events for the accused, and this level of stress can produce deleterious effects on memory retrieval (Deffenbacher, Bornstein, Penrod, & McGorty, 2004). In a study that illustrates this point, Morgan et al. (2004) randomly assigned trainees undergoing

military survival training to endure a realistic high-stress or low-stress mock interrogation. One day later, many of those in the high-stress condition could not even identify their interrogator from a lineup.

In light of the research literature on memory for conversations, the acceptance in court of note taking in lieu of electronic recording, and the significance of the issue for policy and practice, the present research was designed to assess the memorial accuracy of police reports of suspect interrogations. In an experiment conducted at a large Northeastern police station, a sample of experienced investigators examined a mock crime scene, interrogated two innocent male suspects, and submitted a report on their interrogations. For each investigator, expectations were varied from one session to the next, with one of the two suspects presented via demeanor cues as having behaved suspiciously. Unbeknownst to these investigators, all sessions were covertly audio recorded. By later comparing police reports of the sessions with the actual tapes, we addressed two questions. First, what does audio recording reveal about substantive accuracy of police accounts, as measured by intrusion errors, omission errors, and source monitoring confusion, and what does it reveal about the extent to which suspicion elicits confirmation bias in police reports of interrogations? Second, with regard to the fact finder, how might access to the unabridged verbatim transcripts influence others' perceptions of the interrogations and suspects compared to reading only the police reports?

Phase 1 Method

Participants

Phase 1 participants were 18 investigators of various ranks from a large Northeastern police department and 36 male community members recruited via Craigslist to serve as mock suspects. Police participants were recruited during daily roll calls and later via word-of-mouth once the study was ongoing. As noted in the consent form, the stated purpose of the research was “to study the way police investigators conduct suspect interviews and interrogations, and write informative incident reports on their work.” Data from one police participant were excluded after he expressed suspicion that his session had been recorded; a second police participant failed to submit the required written report following his session. Hence, the final sample consisted of $N = 16$ police participants (eight officers, two detectives, and six sergeants) who conducted and reported on a total of 32 suspect interrogations. All sessions were conducted in a vacant office at the police station.

Police participants, half of whom were female, ranged in age from 26 to 55 ($M = 43.94$, $SD = 7.46$) and had an average of 16.33 years of law enforcement experience ($SD = 6.00$). Half had received formal training in suspect interviewing and interrogation; 62.50% estimated that they had conducted over 100 suspect interviews during their careers (if a range was provided, we converted the estimate to the midpoint of that range; the overall median was 250). All suspects were male and ranged in age from 18 to 62 ($M = 34.91$, $SD = 13.86$). A total of 46.88% had previously been arrested; 25% had been suspect-interviewed by police; 28.13% had been convicted of a crime.

Design

Within each session, one police participant investigated a staged crime scene involving a theft and then interrogated two male suspects, both of whom were factually innocent. After looking at the crime scene but prior to interrogating the first suspect, they were told that the first or second of the two suspects (in counterbalanced order) had acted suspiciously upon learning that he would be questioned about the theft. Phase 1 thus employed a two-group (Suspicion: Present vs. Absent) within-subjects design.

Procedure

Two experimenters were involved in each session--one to meet and instruct the police participant, and the other to separately meet and instruct the two suspects. Each session began with a preexperiment questionnaire, at which time

a theft was said to have occurred. This was followed by a police crime scene investigation; two sequential suspect interrogations; and postinterrogation questionnaires. After each session, the police participant was asked to produce and submit a written report within 48 hours that detailed his or her crime scene investigation and interrogations.

Preexperiment questionnaires and mock theft. Upon arrival, the police participant was escorted to an interview room; the two suspects were escorted to a separate waiting room. After giving informed consent, the police participant indicated his or her age and gender and answered several questions concerning background and training.

After suspect participants gave their informed consent, an experimenter directed but did not accompany them, one-at-a-time, to another room where they were to complete a preexperiment questionnaire. The questionnaire asked suspects to indicate their age and whether they had ever been interviewed by police, arrested, or convicted of a crime. On their way to and from this room, suspects passed an unattended briefcase in the hallway that had been staged to look as though a theft had occurred. The briefcase zipper was left ajar, and on the floor next to it were an open lock (which had ostensibly been removed from the briefcase) and a wallet with no cash inside. Suspects were instructed to take notice of--but not to touch--the briefcase. All suspects, therefore, had exposure to the scene of the mock theft but were factually innocent.

Once both suspects had completed the questionnaire and returned to the waiting room, the experimenter informed them that they were the targets of an investigation into a theft committed at the police station and that they would soon be *233 interviewed by a detective. Suspects were paid \$20 for their participation prior to being interrogated and told that they would receive a \$15 bonus if they convinced the detective of their innocence. If they did not succeed, they were told that they would have to return for a second session to receive the additional payment (in actuality, all suspects were paid the full \$35 after the session).

Crime scene investigation. At that point, the second experimenter informed the police participant that an unknown sum of cash had been stolen from a briefcase in the hallway and that the crime scene had been left exactly as it was found. In the context of a lengthy instruction, police were asked to investigate the crime scene, question two suspects who were apprehended nearby around the time that the theft occurred, and solve the crime. The experimenter explained that both suspects had walked past the briefcase to complete paperwork in a nearby room and were out of view when they did so. We provided no inculpatory evidence against the suspects aside from the fact that the money was missing and both were known to have been alone in the vicinity at the time. Police were assured that no actual crime had taken place and that interrogation must be terminated if a participant suspect wanted to stop (the verbatim instructions are available as online supplemental material).

The police participant was then taken to the crime scene, provided with a notepad and digital camera, and given 5 minutes to investigate the area. Afterward, the experimenter escorted the police participant back to the interrogation room and inquired as to whether he or she wanted hard copies of their crime scene photographs for use during the interrogation. If they did, these images were printed, in color, on 4 x 6 in. photo paper.

Suspicion manipulation. While escorting the police participant back to the interrogation room, the experimenter reiterated that he or she would question two theft suspects. In counterbalanced order, the experimenter added that:

The first [or second] guy you're going to question seemed okay, but the second [or first] guy was acting really strange. He looked nervous: He kept pacing back and forth and wouldn't make eye contact with me while I was talking to him.

Given that a mock crime was said to have been committed, this comment about demeanor was designed to selectively raise each police participant's relative a priori suspicion toward one of the two suspects.

Interrogations and postinterrogation questionnaires. Police were told that they would have 20 to 30 min to interrogate each suspect, after which they were to return the suspect to the waiting room. If an interrogation was still ongoing at 20 min, the experimenter knocked on the door to signal to the police participant that it was time to wrap up the interrogation. Suspects were randomly assigned to be interrogated either first or second. All interrogations were surreptitiously audio recorded through a digital voice recorder hidden among office supplies on a desk. Neither police participants nor suspects were preinformed of the fact that the interrogations would be recorded.

After each interrogation, the suspect returned to the waiting room, and both the police participant and suspect completed a self-report questionnaire concerning their perceptions of the interrogation experience. Police participants completed this questionnaire twice, once after each of their interrogations. After the suspect completed his questionnaire, he was fully debriefed, paid, and dismissed.

Police reports. After completing the second postinterrogation questionnaire, police were instructed on how to prepare and submit Incident Reports of their investigation, which were due within 48 hours. To ensure that investigators took the task seriously, these instructions stipulated that reports should be approximately three to five typed pages in length, single spaced, and should consist of three sections: a summary of their crime scene analysis (one page), an account of the first suspect interrogation (1-2 pages), and an account of the second suspect interrogation (1-2 pages). The instructions also specified the sorts of details that should be included--including key questions asked during each interrogation, suspects' answers to these questions, descriptions of each suspect's demeanor, and any indications or impressions of each suspect's involvement. Upon receipt of these reports, police were debriefed and paid \$100 for their participation.

Dependent Measures

Postinterrogation questionnaires. Immediately after each interrogation, police and suspect participants completed a self-report questionnaire consisting of 12 parallel items that measured their perceptions of the interrogation. First, police indicated whether they believed the suspect to be guilty or innocent and indicated their confidence in this impression on a scale ranging from 1 (*not at all*) to 10 (*very*). These two items were later combined to form a guilt-confidence composite score that could range from -10 (*highly confident guilty judgment*) to +10 (*highly confident innocent judgment*). Police also gave continuous ratings of how credible the suspect's denials were, how knowledgeable the suspect seemed to be about the crime, how cooperative the suspect was, how hard they tried to get the suspect to confess, how friendly they were toward the suspect, how confrontational they were toward the suspect, how anxious the suspect was, and how stressful they thought the interrogation was for the suspect. All ratings were made on a scale that ranged from 1 (*not at all*) to 10 (*very*). Finally, police gave two dichotomous yes/no judgments as to whether the suspect had made any suspicious remarks and/or any admissions of guilt during the interrogation.

Suspects simultaneously answered parallel questions. First, they indicated whether they believed the interrogator would perceive them to be guilty or innocent and rated their confidence in that judgment on a 10-point scale. Once again, these two items were combined to form a composite score that could range from -10 (*highly confident that they would be judged guilty*) to +10 (*highly confident that they would be judged innocent*). Suspects also rated how credible their denials were, how much knowledge they had about the crime, how cooperative they were, how hard the interrogator tried to get them to confess, how friendly the interrogator was, how confrontational the interrogator was, how anxious they were, and how stressful the interrogation was for them. These continuous ratings were made on a scale ranging from 1 (*not at all*) to 10 (*very*). Suspects also gave dichotomous yes/no judgments as to whether they had made any suspicious remarks and/or any admissions of guilt. Two suspects (6.25%) neglected to answer one or more of these items, and thus their data are missing from the relevant analyses.

***234 Interrogation tactics.** Audio recordings of all 32 interrogations were transcribed by a professional transcription service. The transcripts were then coded by two independent coders who were blind to our suspicion manipulation and had prior experience coding interrogation transcripts. On the basis of a previously published factor analysis of self-

reported tactic use in a survey of police (Kassin et al., 2007), which formed the basis of a coding scheme introduced by Kassin et al. (2014), we coded transcripts for the presence of 16 interrogation tactics organized into five categories (see Table 1).

TABLE 1

PRESENCE AND FREQUENCY OF 16 CATEGORIZED INTERROGATION TACTICS ACROSS 32 INTERROGATIONS

INTERROGATION TACTICS	% SESSIONS USED	<i>M (SD)</i> OF TIMES USED
Confrontation	56.25	1.63 (1.98)
Calling the suspect a “liar” or accusing him of “lying”	25.00	.63 (1.43)
Pointing out inconsistencies in the suspect's story	25.00	.53 (1.16)
Directly accusing the suspect of the theft	15.63	.25 (.76)
Making expressions of disbelief toward the suspect	12.50	.19 (.47)
Interrupting the suspect's denials	3.13	.03 (.18)
Maximization	59.38	2.09 (2.61)
Threatening the suspect with consequences	56.25	1.94 (2.48)
Exaggerating the seriousness of the offense	15.63	.16 (.37)
Leniency	62.50	2.66 (3.82)
Stating minimization themes that imply leniency	53.13	1.44 (2.47)
Making an explicit offer of leniency for confession	46.88	1.22 (1.72)
False evidence	84.38	1.56 (1.02)^{a1}
Bluffing about future evidence	84.38	1.47 (.98)

Lying about existing evidence	25.00	.38 (.75)
Miscellaneous tactics		
Establishing rapport via small talk	75.00	N.A.
Pressing the suspect to implicate someone else	75.00	N.A.
Encouraging the suspect to admit other illegal acts	68.75	N.A.
Appealing to the suspect's religion/ conscience	21.88	N.A.
Praising or flattering the suspect	15.63	N.A.

Footnotes

^{a1} These values represent the number of different types of evidence about which police participants bluffed or lied.

For three of the five categories (confrontation, maximization, and leniency), coders counted the number of times each individual tactic was used. For some analyses, these counts were later coded as absent (i.e., a count of zero) or present (i.e., any nonzero count). We also computed the sum of tactic uses within each category, enabling us to create a dichotomous measure of whether or not police used at least one tactic from a given category. For the other two categories (false evidence and miscellaneous), coders merely noted whether each tactic was present or absent.

Confrontation tactics involved calling the suspect a liar (e.g., “You just lied to me,” “You're pretty much lying to an authority figure you know”), pointing out inconsistencies in the suspect's story (e.g., “You're flipping your story,” “How do you know there was a card in the wallet? You said you just looked at it”), outright accusing the suspect of committing the crime (e.g., “You did it,” “Okay, well I think that you took the money”), expressing disbelief in the suspect's story (e.g., “That's hard to believe,” “I can tell you right now just from the evidence that I have that your story is not believable”), and interrupting the suspect's attempts to maintain his innocence (as denoted by midsentence hyphens in the suspect portions of the transcripts).

Maximization tactics include threatening the suspect with negative consequences for not confessing (e.g., “If it comes back to you, you're done,” “If you don't tell me up front and then I have to determine that by evidentiary matter then I slam you in court, okay?”), and exaggerating the seriousness of the offense (e.g., “The amount of cash in the wallet can bring the charges from now a theft to a much higher degree,” “There's going to be an additional charge for hindering prosecution”).

Leniency tactics include developing minimization themes that excuse, justify, or otherwise downplay the crime, statements that may lead people to infer leniency in punishment (e.g., “This is something, it's so minor,” “It doesn't make you a bad person, bro”) and explicit offers of leniency or immunity in exchange for a confession (e.g., “I got a little bit of juice over at the courthouse and if you did do it, if you're honest and up front with me, I can work with you,” “I'll get you probation on this,” “I am offering you a deal here then, you know”).

False evidence tactics include the bluff (the assertion that there is evidence to be harvested without the added claim that it implicates the suspect; e.g., “Understand I am going to check the cameras,” “That little padlock and key that you saw, that has been submitted for DNA analysis”) and specific false claims about the existence of actual evidence (e.g., “I already told you there is one person that ID'd you right?” “We have video of you going in that wallet”).

In addition to these basic Reid-technique approaches, a range of miscellaneous tactics were coded as well--including irrelevant small talk aimed at establishing a rapport (e.g., “What kind of food you like cooking?” “What kind of work do you do?”), asking the suspect to implicate someone else (e.g., “Did you see any of the other guys like [John], his brother, whoever, anybody with money take money out of their pockets, anything?” “What about the, um, the older gentleman, was he referencing anything about money?”), encouraging the suspect to admit to other illegal behaviors (e.g., “Are you on probation or anything?” “Okay have you ever been *235 arrested?”), appealing to the suspect's religion or conscience (e.g., “Some of the items had some sentimental value that's what it is. I just want to get those returned to the owner,” “Are you religious at all? There is definitely right and wrong in the universe”), and the use of flattery (e.g., “You seem like a pretty good dude,” “You've obviously got a great relationship with your son, great relationship with your mom”).

Next we coded for the presence or absence of five discernible *police behaviors* (gathering personal information from suspects, showing crime scene photos to the suspect, asking the suspect to empty his pockets, issuing a *Miranda* warning, and obtaining a *Miranda* waiver; see Table 2). We also coded for the frequency of four discernible *suspect behaviors* (denials of guilt, denials of crime-relevant knowledge, disclosures of crime-relevant knowledge, and self-incriminating admissions; see Table 2).

TABLE 2		
<i>PRESENCE AND FREQUENCY OF FIVE POLICE AND FOUR SUSPECT BEHAVIORS ACROSS 32 INTERROGATIONS</i>		
BEHAVIORS CODED	% SESSIONS IN WHICH BEHAVIOR OCCURRED	<i>M (SD)</i> OF OCCURRENCES
<i>Police behaviors</i>		
Obtaining suspect's personal information	90.63	N.A.
Showing crime scene photos to suspect	37.50	N.A.
Asking suspect to empty his pockets	18.75	N.A.
Issuing a <i>Miranda</i> warning	6.25	N.A.
Obtaining a <i>Miranda</i> waiver	6.25	N.A.
<i>Suspect behaviors</i>		
Disclosing crime-relevant knowledge	96.88	16.00 (9.60)

Denying guilt	90.63	7.09 (5.28)
Denying crime-relevant knowledge	75.00	2.66 (3.31)
Making an incriminating admission	12.50	.31 (1.12)

For tactics and behaviors that were coded merely as present or absent (i.e., false evidence tactics, miscellaneous tactics, and police behaviors), coders exhibited an overall agreement rate of 91.67%, $\kappa = 0.83$ [95% CI: 0.77, 0.89], $p < .001$ (within each category, all κ s $> .69$, p s $< .001$). For those that were coded as frequency counts (i.e., confrontation, maximization, and leniency tactics, and suspect behaviors), the overall intraclass correlation (ICC) for our two raters was near perfect, ICC = .97 [95% CI: .96, .98], $p < .001$ (within each category, all ICCs $> .94$, p s $< .001$). Whatever disagreements or discrepancies that remained between coders were resolved via discussion.

Police reports. Two additional independent coders read and coded police reports for disclosures of these same tactics and behaviors, allowing us to compare the contents of the interrogation transcripts against the corresponding police accounts of those same interrogations. For reports, all tactics and behaviors were coded as either present (i.e., the police participant described the use of a given tactic or the presence of a given behavior) or absent (i.e., the given tactic or behavior was not noted in the police report). These coders exhibited an overall agreement rate of 97.10%, $\kappa = 0.89$ [95% CI: 0.85, 0.93], $p < .001$ (within each category, all κ s $> .65$, p s $< .001$). Once again, disagreements were resolved by discussion.

Phase 1 Results

Length of Interrogations

On average, interrogations lasted for 16.41 min ($SD = 5.50$; Range: 3.32 - 26.17) and contained 2,438.31 words ($SD = 984.11$; Range: 866 - 4,481). Our suspicion manipulation had no effect on the duration, $t(15) = 0.29$, $p = .780$, $d = 0.07$ [95% CI: -0.49, 0.63], or word count, $t(15) = 0.51$, $p = .619$, $d = 0.13$ [95% CI: -0.43, 0.68], of interrogations (note that all CIs reported here and thereafter pertain to effect sizes). On average, interrogations of the second suspect were longer in word count, $t(15) = 2.22$, $p = .042$, $d = 0.55$ [95% CI: 0.00, 1.11], but not in duration, $t(15) = 1.64$, $p = .121$, $d = 0.41$ [95% CI: -0.15, 0.97]. On average, 56.60% of all words were spoken by the interrogator ($SD = 11.63\%$; Range: 28.05-74.66%), and neither Suspicion, $t(15) = 0.56$, $p = .586$, $d = 0.14$ [95% CI: -0.42, 0.70], nor Order, $t(15) = 0.87$, $p = .401$, $d = 0.22$ [95% CI: -0.34, 0.77] affected this percentage.

No significant order effects were found on any other dependent measures, including responses to the postinterrogation questionnaires and the presence and frequency of all coded tactics and behaviors, with one exception: Suspects who were interrogated second offered more denials of crime-relevant knowledge ($M = 3.81$, $SD = 4.04$) than those who were interrogated first ($M = 1.50$, $SD = 1.86$), $t(15) = 2.21$, $p = .043$, $d = 0.55$ [95% CI: -0.01, 1.11]. Consequently, all data were collapsed across order for all subsequent analyses.

Postinterrogation Questionnaire

A 2 (Suspicion: Present vs. Absent) x 2 (Source: Police vs. Suspect) repeated-measures MANOVA was performed on the nine continuous items on the postinterrogation questionnaire (including guilt-confidence composite scores) to test for discrepancies between suspect and police perceptions of the interrogation. A multivariate effect of Source emerged, Wilks' $\Lambda = .10$, $F(9, 5) = 5.16$, $p = .043$, with significant differences on four items. Neither the multivariate effect of Suspicion, Wilks' $\Lambda = .18$, $F(9, 5) = 2.46$, $p = .167$, nor the Source x Suspicion interaction, Wilks' $\Lambda = .23$, $F(9, 5) = 1.86$, $p = .257$, was significant.

Follow-up univariate ANOVAs were performed on the four items that differed as a function of Source (see Figure 1). These analyses indicated that suspects were more confident that they would be judged as innocent ($M = 4.07$, $SD = 7.04$) than police were confident in their innocence ($M = 0.54$, $SD = 6.95$), $F(1, 13) = 4.91$, $p = .045$, $d = 0.35$ [95% CI: -0.07, 0.77]; suspects rated their denials as more credible ($M = 8.25$, $SD = 1.80$) than police rated their denials ($M = 5.61$, $SD = 2.17$), $F(1, 13) = 26.92$, $p < .001$, $d = 0.88$ [95% CI: 0.46, 1.30]; and suspects rated themselves as more cooperative ($M = 9.11$, $SD = 1.37$) than police rated them ($M = 6.89$, $SD = 2.48$), $F(1, 13) = 24.21$, $p < .001$, $d = 0.84$ [95% CI: 0.42, 1.26]. In addition, police rated the interrogations as more stressful for suspects ($M = 5.07$, $SD = 2.28$) than suspects rated it for themselves ($M = 3.46$, $SD = 2.15$), $F(1, 13) = 7.57$, $p = .016$, $d = 0.63$ [95% CI: 0.21, 1.05]. The remaining four items did not differ by Source. Police and suspects, respectively, rated police as friendly ($M_s = 6.93$ & 6.86 , $SD_s = 1.88$ and 2.19) and moderately *236 confrontational ($M_s = 4.86$ and 4.32 , $SD_s = 2.26$ and 2.60), and suspects as moderately anxious ($M_s = 5.18$ and 4.32 , $SD_s = 2.28$ and 2.36) and knowledgeable about the crime ($M_s = 4.82$ and 5.61 , $SD_s = 2.79$ and 2.86).

TABULAR OR GRAPHIC MATERIAL SET FORTH AT THIS POINT IS NOT DISPLAYABLE

Figure 1. Effects of source (suspect vs. police) on four (of nine) continuous self-report items (Phase 1). Note. Higher scores on the composite measure indicate greater confidence in the suspect's projected innocence.

We separately analyzed three dichotomous items on the postinterrogation questionnaire. First, police judged 13 of the 32 suspects as guilty (40.62%). As with other dependent measures, suspicion did not affect these judgments, McNemar $X^2(1) = 0.82$, $p = .549$, nor did it affect suspects' predictions of whether they would be judged guilty, McNemar $X^2(1) = 0.11$, $p = 1.00$. Police judgments of the suspect's guilt were unrelated to suspects' predictions of whether they would be judged as guilty, McNemar $X^2(1) = 0.60$, $p = .607$. Second, police indicated that 13 of the 32 suspects (40.62%) made suspicious remarks during the interrogation. Once again, suspicion did not affect police responses to this item, McNemar $X^2(1) = 0.14$, $p = 1.00$ (note that three "suspicious remarks" suspects were judged innocent; three others not seen as having made suspicious remarks were judged guilty.). Police and suspect responses to this item were also unrelated to each other, McNemar $X^2(1) = 0.25$, $p = .804$. Third, only one suspect was judged by police as having made an admission of guilt; no suspects reported having admitted guilt.

Interrogation Tactics and Behaviors

Frequencies and descriptive statistics for all 16 coded interrogation tactics and five tactic categories are shown in Table 1. Out of 32 interrogations, 56.25% featured the use of one or more confrontation tactics, 59.38% featured one or more maximization tactics, 62.50% featured one or more leniency tactics, and 84.38% featured one or more false evidence tactics. The average interrogation featured 1.63 uses of confrontation ($SD = 1.98$), 2.09 uses of maximization ($SD = 2.61$), and 2.66 uses of leniency ($SD = 3.82$).

With respect to presentations of false evidence, 84.38% of interrogations featured bluffs about evidence, and the average interrogation included bluffs about 1.47 different types of evidence ($SD = 0.98$). Police most often bluffed about having surveillance footage (75% of interrogations), followed by fingerprints (46.88%), DNA (18.75%), and eyewitnesses (3.13%). One quarter of interrogations featured outright lies about existing evidence, and the average interrogation included lies about 0.38 different types of evidence ($SD = 0.75$). Police most often lied about incriminating surveillance footage (21.88% of interrogations), followed by DNA (9.38%), fingerprints (3.13%), and eyewitnesses (3.13%).

Descriptive statistics for the five coded police behaviors and four suspect behaviors are shown in Table 2. Most police solicited personal information from the suspect (90.63%); fewer showed crime scene photos (37.50%), asked the suspect to empty his pockets (18.75%), or read *Miranda* warnings (6.25%). Virtually all suspects (96.88%) disclosed some crime-relevant knowledge during the interrogation. On at least one occasion, most suspects denied guilt (90.63%) and denied crime-relevant knowledge (75.00%); relatively few suspects made incriminating statements (12.50%).

Suspicion Manipulation

Suspicion did not affect any of the coded tactics, namely, whether the interrogator exhibited one or more uses of confrontation, maximization, leniency, or false evidence, all McNemar X^2 s ≤ 1.00 , $ps \geq .60$; the number of uses of confrontation, maximization, or leniency tactics, all $ts < 1$, $ps > .75$, $ds < 0.10$; or the number of different bluffs, $t(15) = 0.59$, $p = .566$, $d = 0.15$ [95% CI: -0.41, 0.70], or lies, $t(15) = 0.19$, $p = .849$, $d = 0.05$ [95% CI: -0.51, 0.61]. Suspicion likewise did not influence the likelihood of any of the five coded police behaviors, all McNemar $X^2 \leq 2.00$, $ps \geq .50$, the likelihood of any of the four coded suspect behaviors, all McNemar $X^2 \leq 0.50$, $ps > .70$, or the number of times that any of the four suspect behaviors occurred, all $ts \leq 1.05$, $ps > .30$, $ds \leq 0.15$.

Comparison of Transcripts and Reports

In comparison to the interrogation transcripts, which contained an average of 2,438.31 words ($SD = 984.11$; Range: 866 - 4,481), police reports contained an average of 1,224.44 words ($SD = 365.52$; Range: 570 - 1,875). Descriptions of the crime scene contained fewer words ($M = 258.06$, $SD = 76.31$) than descriptions of either the first ($M = 452.44$, $SD = 174.94$) or the second ($M = 478.31$, $SD = 183.06$) interrogation, $F(2, 30) = 23.01$, $p < .001$, $\eta^2_p = .61$, which did not differ in word count.

We sought to measure both the accuracy and completeness of police reports by noting errors of commission and omission in relation to the coded transcripts. For this purpose, we compared the *237 presence of all coded tactics and behaviors in the interrogation transcripts against the presence of those same tactics and behaviors in the corresponding police reports. Errors of commission were extremely rare. Across 25 coded tactics and behaviors in 32 interrogations (a total of 800 coded details), only four errors of commission were found (0.50%). One interrogator mistakenly reported having lied about evidence to one suspect and having shown crime scene photos to the other; a second interrogator mistakenly reported using minimization tactics; and a third mistakenly reported having offered leniency in exchange for a confession.

In contrast, errors of omission were prevalent. Figure 2 shows the frequencies of usage and reporting for all coded tactics and behaviors. As noted earlier, a majority of interrogations included the use of one or more tactics involving confrontation (56.25%), maximization (59.38%), leniency (62.50%), and false evidence (84.38%). When used, however, the corresponding reports described one or more of these tactics only 22.22%, 15.79%, 40%, and 66.67% of the time, respectively (e.g., in 77.78% of reports of interrogations in which confrontation was used one or more times, no mention of any confrontation tactic was made). In short, a number of different tactics were frequently used during interrogations but did not appear in corresponding police reports.

TABULAR OR GRAPHIC MATERIAL SET FORTH AT THIS POINT IS NOT DISPLAYABLE

Figure 2. Occurrence and reporting of interrogation tactics, police behaviors, and suspect behaviors across 32 interrogations and corresponding police reports.

With respect to suspect behaviors, virtually all suspects (96.88%) disclosed some crime-relevant knowledge regarding the mock theft, and these disclosures were almost always reported by police (96.77%). Most suspects also issued denials of guilt (90.63%), which were also typically reported (89.66%). However, while 75% of suspects denied crime-relevant knowledge at least once, only 33.33% of the corresponding reports noted such denials (in contrast, three out of four suspects who made incriminating admissions were reported to have done so).

To sum up: Two sets of findings emerged from Phase 1 of our study. First, police and suspects diverged in their perceptions of the interrogations, with suspects seeing themselves as more credible in their denials, more cooperative,

and more confident in their presentation of innocence. Second, a comparison of interrogation transcripts and police reports indicated that errors of omission were prevalent, with police underreporting the use of various tactics. In light of these results, Phase 2 assessed whether observers' impressions of the interrogations were differently influenced *238 by the police reports that were generated relative to full and objective records.

Phase 2 Method

Participants and Design

Participants in Phase 2 were 96 undergraduates, 64 of whom were female, with a mean age of 19.45 ($SD = 2.84$). A total of 38.54% self-identified as Hispanic, 17.71% as White, 15.63% as Asian, 12.50% as Black, and 15.63% as multiracial or Other.

To determine whether outside fact finders would form different or more accurate impressions when they had access to verbatim transcripts of suspect interrogations, compared to secondhand police reports, participants read a brief summary of the case and were randomly assigned to either read a police report or an interrogation transcript of one of the 32 Phase 1 interrogations. These conditions enabled us to compare directly the two alternative written accounts of the process. Next we sought to assess what impressions participants would form if they also listened to audiotapes of these same sessions, thereby having access to vocal and paralinguistic cues emanating from both police and suspect participants (e.g., prosodics such as pitch, volume, stress, tempo, pitch, rhythm, pauses, points of stress, and intonation--the kinds of cues that often lead laypeople to draw erroneous inferences of truth and deception; see DePaulo et al., 2003; Bond & DePaulo, 2006). Thus, a third group of participants--equipped with headphones and an MP3 player--listened to the audio recording while following along on the transcript. This group enabled us to compare the two verbatim conditions--one written, the second accompanied by audio cues. Because no significant effects of manipulated suspicion were obtained in Phase 1, this second experiment combined rather than distinguished these two groups of suspects in all analyses.

Procedure

Participants completed the study in sessions that included from four to eight individuals. All participants in a given session were in the same experimental condition, but they read and/or listened via headphones to different Phase 1 interrogations. After giving their informed consent and providing basic demographic information (i.e., age, gender, and race), all participants read the same one-page background summary of the crime and investigation that occurred during Phase 1. The summary described the theft of an unknown amount of cash from a briefcase inside a police station, described the crime scene and the contents of the briefcase, and explained that two suspects were identified who were known to have been in the vicinity of the briefcase around the time of the theft. Participants were told that they would now be given an account of the interrogation of one of these two suspects. Participants then read the police report (*report* condition) or the full transcript (*transcript* condition), or they listened to an audio recording (*audio* condition) of one of the 32 Phase 1 interrogations. When they were finished, they completed a questionnaire that measured their perceptions of the interrogator and suspect, after which they were fully debriefed.

Dependent Measures

The questionnaire administered to Phase 2 participants contained 13 items that paralleled those answered by police and suspects in Phase 1. First, participants indicated whether they believed the suspect to be guilty or innocent and rated their confidence in that judgment on a scale from 1 (*not at all*) to 10 (*very*). As in Phase 1, these items were combined to form a guilt-confidence composite score that could range from -10 (*highly confident guilty judgment*) to +10 (*highly confident innocent judgment*). Participants also gave continuous ratings of how credible the suspects' denials were, how knowledgeable the suspect was about the crime, how cooperative the suspect was, how hard the interrogator tried to get the suspect to confess, how friendly the interrogator was, how confrontational the interrogator was, how anxious

the suspect was, and how stressful they thought the interrogation was for the suspect. One new continuous item was added for Phase 2, which asked participants to rate how much pressure the interrogator placed on the suspect during the interrogation. These nine continuous ratings were each given on a scale from 1 (*not at all*) to 10 (*very*). Participants also gave dichotomous yes/no judgments as to whether the suspect had made any suspicious remarks and/or any admissions of guilt during the interrogation. Four participants (4.17%) neglected to answer one or more items, so their data are missing from the relevant analyses.

Phase 2 Results

Two a priori sets of comparisons framed our analyses. First and foremost, we sought to compare the two written accounts of the interrogations--secondhand police reports versus verbatim transcripts. Next we compared the two verbatim accounts--written transcripts alone versus accompanied by an audiotape.

Written Police Reports Versus Transcripts

Our most important prediction was that reading transcripts of interrogations would improve fact-finding accuracy relative to reading the accounts contained in police incident reports by rendering participants less likely to judge innocent suspects as guilty. Supporting this prediction, participants in the Transcript condition misjudged the suspect as guilty less often than did those in the Report condition (9.38% vs. 31.25%, respectively), $X^2(1) = 4.73, p = .030, \# = .27, OR = 4.39$ [95% CI: 1.08, 17.86].

We then compared the Transcript and Report conditions in terms of their continuous ratings of the interrogator and suspect on our questionnaire. A one-way MANOVA on these 10 items did not reach a conventional level of significance, Wilks' $\# .74, F(10, 47) = 1.65, p = .123$, but it did reveal a large effect size ($\#^2_p = .26$) and suboptimal power ($1-\# = .71$). In light of this, and to more precisely focus on our most pertinent system-relevant measures, we performed univariate t tests designed to directly test our prediction that presentation medium would in particular impact perceptions of *guilt* and *coercion*.

Significant univariate differences were found on three items, each of which showed a medium to large effect size (see Figure 3; Cohen, 1988). Consistent with the aforementioned binary judgments, participants who read a transcript were more confident in the suspect's innocence ($M = 5.86, SD = 4.82$) than were those who read the corresponding police report ($M = 2.34, SD = 7.12$), $t(56) = 2.20, p = .032, d = 0.59$ [95% CI: -0.95, 2.14]. Importantly as well, compared to those in the Report condition, participants who read a transcript also rated the interrogator as having ***239** exerted more pressure on the suspect ($M_s = 6.93$ and $5.59, SD_s = 2.36$ and 2.43), $t(56) = 2.14, p = .037, d = 0.57$ [95% CI: 0.03, 1.18], and as having tried harder to obtain a confession ($M_s = 7.59$ and $5.79, SD_s = 2.50$ and 2.70), $t(56) = 2.62, p = .011, d = 0.70$ [95% CI: 0.04, 1.36].

TABULAR OR GRAPHIC MATERIAL SET FORTH AT THIS POINT IS NOT DISPLAYABLE

Figure 3. Comparison of written transcripts and reports on Phase 2 perceptions of interrogators and suspects. Note. Higher scores on the composite measure indicate greater confidence in the suspect's innocence.

No differences were found on the other continuous ratings, all $t_s < 1.15, p_s > .26$. Overall, participants in the Transcript and Report conditions, respectively, rated the interrogation as moderately stressful ($M_s = 5.28$ and $5.66, SD_s = 2.83$ and 3.11) and the suspect as moderately anxious ($M_s = 5.24$ and $5.48, SD_s = 2.69$ and 2.89), credible ($M_s = 6.72$ and $6.21, SD_s = 1.79$ and 2.32), and knowledgeable ($M_s = 5.48$ and $5.38, SD_s = 2.52$ and 2.62)--and as highly cooperative ($M_s = 8.52$ and $8.14, SD_s = 1.38$ and 2.01). Transcript and Report participants also rated the interrogator both as somewhat friendly ($M_s = 6.07$ and $5.31, SD_s = 2.88$ and 2.27) and somewhat confrontational ($M_s = 6.07$ and $5.41, SD_s = 2.30$ and 2.15).

Two additional questionnaire items asked whether suspects had made any suspicious remarks or admissions of guilt. Participants in the Transcript condition were no more or less likely than those in the Report condition to believe that the suspect had made suspicious remarks (37.50% vs. 28.13%, respectively), $X^2(1) = 0.64, p = .424, \# = .10, OR = 1.53$ [95% CI: 0.54, 4.39], nor were they more or less likely to believe that he had admitted guilt (13.79% vs. 6.90%, respectively), $X^2(1) = 0.74, p = .389, \# = .11, OR = 2.16$ [95% CI: 0.36, 12.84].

Verbatim Transcripts Versus Audiotapes

Next, we compared the judgments of participants in the Transcript and Audio conditions who received the same verbal content but presented in a different modality (in writing vs. audio recording). Interestingly, we found that participants in the Audio condition misjudged the suspect as guilty more often than did those in the Transcript condition (37.50% vs. 9.38%, respectively), $X^2(1) = 7.05, p = .008, \# = .33, OR = 5.80$ [95% CI: 1.45, 23.23], a difference that is not surprising in light of research showing that people are not intuitively accurate judges of truth and deception. Audio participants were not, however, more likely to believe that their suspect had made suspicious remarks (53.13% vs. 37.50%, respectively), $X^2(1) = 1.58, p = .209, \# = .16, OR = 1.89$ [95% CI: 0.70, 5.12], or that he had made an admission of guilt (12.90% in both conditions), $X^2(1) = 0.00, p = 1.00, \# = .00, OR = 1.00$ [95% CI: 0.23, 4.42].

A one-way MANOVA on continuous ratings of the interrogator and suspect did not reach significance, Wilks' $\# = .73, F(10, 51) = 1.85, p = .075$. Once again, however, this analysis revealed a large effect size ($\#^2_p = .27$) and a relative lack of power ($1-\# = .78$). Follow-up t tests on the most relevant continuous items revealed two significant univariate differences with medium to large effect sizes. First, participants in the Audio condition were less confident in the suspect's innocence ($M = 1.19, SD = 7.55$) than were those in the Transcript condition ($M = 5.97, SD = 4.68$), $t(60) = 2.99, p = .004, d = 0.77$ [95% CI: -0.77, 2.31]. Second, participants in the Audio condition rated the interrogator as having exerted less pressure on the suspect ($M = 5.39, SD = 3.13$) than did those in the Transcript condition ($M = 6.89, SD = 2.46$), $t(60) = 2.10, p = .040, d = 0.54$ [95% CI: -0.15, 1.23]. No significant differences were found on the remaining items, all $ts < 1.87, ps > .06$.

Discussion

In light of the problems associated with confession evidence, numerous social scientists, legal scholars, and practitioners have recommended a policy reform that would require the electronic recording of entire suspect interviews and interrogations--not just the resulting confessions (e.g., American Psychology-Law Society white paper--Kassin et al., 2010). Historically, such a requirement has proved controversial, drawing opponents from the law enforcement community who have speculated about the possible adverse effects on police, suspects, and juries (e.g., Inbau et al., 2001; for an overview of opposition arguments see Sullivan, 2008; Thurlow, 2005).

The present research was designed to test for an important possible benefit of audio recording full interrogations: that these recordings provide a more accurate factual account of the interrogation behavior of police and suspects than would otherwise be derived from police reports. Indeed, a common argument for recording interrogations is that it is the most effective way to memorialize the process by which a statement was taken and, hence, increase the fact-finding accuracy of judges and juries. In current practice, whereby detectives produce reports from contemporaneous or retrospective notes, disputes often arise as to a number of issues concerning the behavior of police and suspects--most notably, whether the police Mirandized suspects in a timely manner, whether they used certain coercive tactics, and whether *240 the details contained within a confession originated with the suspect or came about through a process of contamination.

Police notes from interviews and interrogations are routinely accepted in the courts in lieu of an objective recording. Yet the accuracy and impact of these notes has never been tested in the context of a suspect interrogation. In light of

the importance of the policy in question as well as basic research on memory for conversations, we conducted a two-phased experiment. In Phase 1, experienced police officers investigated a mock crime scene, interviewed two innocent suspects--one of whom was presented as suspicious on the basis of his behavior--and then filed an incident report. All sessions were covertly audio recorded; these recordings were later used to assess the accuracy of the reports. In Phase 2, lay participants read either the police report or the interrogation transcript, or they listened to an audiotape. Our goal was to compare the accuracy of these "fact finders" as a function of the information they were provided.

In Phase 1, two notable sets of results were obtained. First, in parallel postinterrogation questionnaires, police and their suspects diverged in their perceptions of the suspect's behavior. Specifically, suspects believed that they would be seen as more innocent than the police actually believed them to be; they also rated their denials as more credible, their stress levels as lower, and their behavior as more cooperative. Along with the finding that police judged 41% of suspects as guilty, and has having made suspicious remarks, this result demonstrates a phenomenon often seen in real trials: That detectives and defendants often testify disparately as to what transpired during the process of interrogation.

The second key result emerged from the comparison between the tactics that police actually used during their interrogations, as later coded from the tapes, and the tactics police said they used in their incident reports. Consistent with past research using a mock crime paradigm (e.g., Kassin et al., 2014), in naturalistic field settings involving real interrogations (e.g., Feld, 2013; Leo, 1996), and in self-report surveys of police (e.g., Kassin et al., 2007), we found that police participants commonly used confrontation, maximization, minimization and leniency, and presentations of false evidence in the form of bluffing and outright lies--even in the context of interrogations that were brief and without high-stakes consequences.

In this regard, two important points are worth noting. First, closely replicating the mock interrogation results reported by Kassin et al. (2014), the police participants in our study used the false evidence ploy at a very high rate. This result contrasts with more realistic field-based data (e.g., Leo, 1996) and self-report surveys (e.g., Kassin et al., 2007) indicating that this tactic is used more sparingly. Second, when the usage frequencies were compared with the incident reports, even though police participants were instructed to state as close to verbatim as possible what tactics they used, numerous errors of omission were observed. Specifically, whereas most interrogations involved the use of confrontation, maximization, leniency, and false evidence at rates of 56%, 59%, 63%, and 84%, respectively, corresponding reports described these tactics only 22%, 16%, 40%, and 67% of the time, respectively. In actual cases, such underreporting, whether purposeful or inadvertent, could influence judges who rule on the voluntariness of statements taken and juries who rule on their credibility.

In Phase 2, we sought primarily to determine whether outside fact finders would form different or more accurate impressions when they had access to written transcripts of suspect interviews compared to the written police reports. On the most important dependent measure, perceptions of the suspect's guilt or innocence, the results confirmed expectations: Participants were significantly less likely to misjudge the innocent suspect as guilty when they read a written transcript of the interrogation than when they read the corresponding police report (9% vs. 31%). Moreover, participants who read a police report, compared to those who read the full transcript, believed that police officers had applied less pressure on the suspect and tried less hard to get the suspect to confess. In short, observers whose information basis was one of the police reports relative to one of the transcripts, saw the process as less pressure-filled and the innocent suspect as more guilty.

In addition to comparing the two alternative forms of written information-- police reports and transcripts--we wondered what impressions observers would have if they listened to an interrogation audiotape. On the one hand, this richer medium of presentation presents exactly the same verbal content along with the transcripts--perhaps, therefore, with the same result. On the other hand, audiotapes add complex vocal and paralinguistic cues from both police and suspect participants--the kinds of cues that often lead laypeople to draw unwarranted and erroneous inferences of truth and deception (e.g., DePaulo et al., 2003; Bond & DePaulo, 2006). Suggesting that the latter cues influenced perceptions,

apart from verbal content, several results showed that participants who heard an audiotape, relative to those who merely read a transcript, fell prey to some of these effects. Most notably, for example, 38% misjudged the suspect to be guilty, a number that was significantly higher than in the transcript-only condition.

As a matter of policy, one could argue from our results that perhaps strictly content-focused interrogation transcripts--without access to audio and visual cues--would provide a sufficient basis for fact finding. Such a conclusion would not be warranted. While a transcript fully communicates the verbal text of a police--suspect interaction, it does not depict potentially important aspects of the suspect (e.g., his or her physical condition, appearance, attire, voice, and demeanor; whether he or she is seated in a corner or handcuffed) or the police officers (e.g., their number, size, and proximity to the suspect; whether they are uniformed or in plain clothes; whether weapons are visible; whether they raise their voices). In this regard, extensive research indicates that fact finders render more balanced and accurate judgments from "equal focus" video recordings that show both the suspect and police rather than one or the other (Lassiter, Diamond, Schmidt, & Elek, 2007; Lassiter, Geers, Handley, Weiland, & Munhall, 2002).

Summary, Implications, and Limitations

The most important signal to emerge from our study is one that strongly supports a requirement that all suspect interviews and interrogations be recorded and transcribed in order to provide a more accurate account of the process and improve the fact-finding performance of judges and juries. This signal is embodied in the following main findings: (a) Police and suspects diverged in their perceptions of the suspect's behavior during the interrogation sessions in which both parties were present; (b) Police committed frequent errors of omission in their Incident Reports, underreporting their use of confrontation, maximization, leniency, and presentations of false evidence; and (c) Phase 2 participants who read a police report, compared to those who read a full verbatim *241 transcript, perceived the process as less pressure-filled and were more likely to misjudge innocent suspects as guilty.

Taken as a whole, these findings help to explain the second problem with false confessions (the first being that they occur): That they are too often believed by judges and juries equipped only with secondhand information about the process by which the statements were elicited (Kassin, 2012). On this issue, results from Phase 2 suggest two important points. First, participant observers were aided by having access to the more accurate content provided by verbatim transcripts, suggesting, perhaps, this previously neglected potential benefit of electronic recording. Second, participant observers did not benefit from the addition of audio recordings of the interrogations, leaving open the empirical question, untested in our study, of whether additional access to equal-focus video would improve fact-finding performance (for a discussion, see Snyder, Lassiter, Lindberg, & Pinegar, 2009).

One might wonder whether the errors of omission we observed, in the form of underreporting the use of interrogation tactics, were purposeful or inadvertent. As a strictly empirical matter, our data do not permit us to weigh in on this issue. However, given the context (i.e., these errors were observed in a mock-crime-and-investigation study; police participants wrote Incident Reports that would not later be shown to a prosecutor, entered into evidence at trial, or become the subject of sworn testimony), common sense would suggest that these errors were inadvertent. Such inadvertence is often seen in the contamination of false confessions with accurate crime details, a phenomenon observed in actual cases (to illustrate, see Trainum, 2007; for reviews, see Garrett, 2010, 2015; also see Lamb et al., 2000). In this regard, it is important to note that we did not test for--and our estimates do not account for the possibility of--conscious or motivated failures of police to recall or report certain aspects of the interrogations they conduct. Because we asked for all reports to be submitted within 48 hours, we also did not test for the possibility that reporting errors would increase as a function of longer intervals between interrogations and reports.

The present research is potentially limited in important ways. Our police participants were trained and seasoned professionals, ranging in age from 26 to 55 and having an average of 16 years of law enforcement experience, which included numerous suspect interviews. In the context of our mock-crime-and-investigation paradigm, however, these

participants were limited to 5 minutes for crime scene analysis and two relatively brief interrogations--each lasting an average of only 17 min, and a 48-hr time limit for submission of reports. Whether our results would generalize to longer and more consequential investigations remains an important empirical question for further research. On the one hand, one might argue that our police participants were not as motivated to recall the details of their mock interrogations as they are in actual practice, causing us to overestimate errors of omission and their effects. On the other hand, one might argue from a cognitive perspective that because it is so much easier to recall brief conversations that last only a few minutes, rather than hours-long interrogations, our results underestimate the potential for errors of omission and do not adequately test for possible errors of commission, including source attribution errors.

Our results are also limited by the fact that our manipulation of prior suspicion had no significant effects whatsoever--not on the length of interrogations in minutes or words, participants' perceptions of the exchange, police perceptions of the suspect's guilt, or the coded interrogation tactics that were used. We had hoped to assess whether the accuracy of police reports and the impressions they elicited in observers were moderated by investigators' preexisting degree of suspicion. In prior research, suspicion was varied by the presentation of a base rate (i.e., 80 vs. 20% of suspects in this study are guilty of the mock crime; see Kassin, Goldstein, & Savitsky, 2003). In this study, however, because each investigator was set to conduct two interrogations, we manipulated relative suspiciousness by describing the demeanor of the two suspects--one as calm, the other as anxious and evasive. Lacking a manipulation check to ensure that investigators drew the intended inferences from the experimenter's description of the two suspects, we cannot adequately evaluate the possible effects of suspicion on accuracy and bias in police reports. Testing this confirmation bias hypothesis thus remains an important avenue for follow-up research.

We should also comment on the mock-crime-and-investigation paradigm we used, modeled after that previously reported in Kassin et al. (2014). Conducted in a police station and involving a collaboration of experienced law enforcement participants, this experiment contained a high level of ecological validity. As noted by Kassin et al. (2014), however, such data are difficult to collect during the workday from on-duty and off-duty officers, detectives, and sergeants. For that reason, our sample was smaller than we had hoped it would be, thereby limiting the power of our study and hence our ability to analyze for individual differences in experience or training among police participants.

One final limitation concerns the medium through which Phase 2 participants, mimicking fact finders, made their judgments. Specifically, they read a brief description of the crime followed by the "raw data" of a police report or verbatim transcript or audio recording of an interrogation. Police participants who had interrogated mock suspects did not testify as to their experience--and they were not cross examined. Although research suggests that the process of cross examination may have variable effects--for example, helping jurors to become more discerning of scientific experts (Austin & Kovera, 2015), yet impairing an eyewitness's memory (Valentine & Maras, 2011)--it is nevertheless the natural process through which fact finders are informed about interviews and interrogations. More research is needed to determine if cross-examination serves to correct for the underreporting of tactics used.

With an accumulation of DNA exonerations illuminating the problem of false confessions, and with research indicating the dual risk that innocent people might confess to crimes they did not commit and that judges and juries may well believe these false confessions, it is easy to understand calls to reform that focus on the recording of interrogations. Limitations notwithstanding, the present study adds to a growing empirical literature indicating the need for such a requirement, if only to ensure the accuracy and completeness of memorial accounts of key transactions between police and their suspects--accounts that form the basis of decisions routinely made by judges and juries.

References

American Bar Association. (2004). *Resolution 8A--Videotaping custodial interrogations*. Approved February, 9, 2004.

American Psychological Association. (2014). *Resolution on interrogations of criminal suspects*. Retrieved from <http://www.apa.org/about/policy/interrogations.aspx>

*242 Appleby, S. C., Hasel, L. E., & Kassin, S. M. (2013). Police-induced confessions: An empirical analysis of their content and impact. *Psychology, Crime & Law, 19*, 111-128. <http://dx.doi.org/10.1080/1068316X.2011.613389>

Austin, J. L., & Kovera, M. B. (2015). Cross-examination educates jurors about missing control groups in scientific evidence. *Psychology, Public Policy, and Law, 21*, 252-264. <http://dx.doi.org/10.1037/law0000049>

Boetig, B. P., Vinson, D. M., & Weidel, B. R. (2006). Revealing incommunicado. *FBI Law Enforcement Bulletin, 75*, 1-8.

Bond, C. F., Jr., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10*, 214-234. http://dx.doi.org/10.1207/s15327957pspr1003_2

Bruck, M., Ceci, S. J., & Francoeur, E. (1999). Accuracy of mothers' memories of conversations with their preschool children. *Journal of Experimental Psychology: Applied, 5*, 89 -106. <http://dx.doi.org/10.1037/1076-898X.5.1.89>

Buckley, D. M., & Jayne, B. C. (2005). *Electronic recording of interrogations*. Eagle River, WI: Hahn Printing.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Deffenbacher, K. A., Bornstein, B. H., Penrod, S. D., & McGorty, E. K. (2004). A meta-analytic review of the effects of high stress on eyewitness memory. *Law and Human Behavior, 28*, 687-706. <http://dx.doi.org/10.1007/s10979-004-0565-x>

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129*, 74-118. <http://dx.doi.org/10.1037/0033-2909.129.1.74>

Drizin, S. A., & Reich, M. J. (2004). Heeding the lessons of history: The need for mandatory recording of police interrogations to accurately assess the reliability and voluntariness of confessions. *Drake Law Review*, 52, 619-646.

Federal Rules of Evidence. (2015). Arlington, VA: Federal Evidence Review. Retrieved from <http://federalevidence.com/downloads/rules.of.evidence.pdf>

Feld, B. C. (2013). *Kids, cops, and confessions: Inside the interrogation room*. New York, NY: New York University Press.

Garrett, B. L. (2010). The substance of false confessions. *Stanford Law Review*, 62, 1051-1119.

Garrett, B. L. (2015). Contaminated confessions revisited. *Virginia Law Review*, 101, 395-453.

Geller, W. A. (1993). *Videotaping interrogations and confessions: A preliminary examination of issues*. Washington, DC: National Institute of Justice.

Gudjonsson, G. H. (2003). *The science of interrogations and confessions: A handbook*. Chichester, UK: Wiley.

Inbau, F., Reid, J., Buckley, J. P., & Jayne, B. C. (2001). *Criminal interrogation and confessions* (4th ed.). Gaithersburg, MD: Aspen.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, 114, 3-28. <http://dx.doi.org/10.1037/0033-2909.114.1.3>

Kassin, S. M. (1997). The psychology of confession evidence. *American Psychologist*, 52, 221-233. <http://dx.doi.org/10.1037/0003-066X.52.3.221>

Kassin, S. M. (2005). On the psychology of confessions: Does innocence put innocents at risk? *American Psychologist*, 60, 215-228. <http://dx.doi.org/10.1037/0003-066X.60.3.215>

Kassin, S. M. (2012). Why confessions trump innocence. *American Psychologist*, *67*, 431-445. <http://dx.doi.org/10.1037/a0028212>

Kassin, S. M., Drizin, S. A., Grisso, T., Gudjonsson, G. H., Leo, R. A., & Redlich, A. D. (2010). Police-induced confessions: Risk factors and recommendations. *Law and Human Behavior*, *34*, 3-38. <http://dx.doi.org/10.1007/s10979-009-9188-6>

Kassin, S. M., Goldstein, C. C., & Savitsky, K. (2003). Behavioral confirmation in the interrogation room: On the dangers of presuming guilt. *Law and Human Behavior*, *27*, 187-203. <http://dx.doi.org/10.1023/A:1022599230598>

Kassin, S. M., & Gudjonsson, G. H. (2004). The psychology of confession evidence: A review of the literature and issues. *Psychological Science in the Public Interest*, *5*, 33-67. <http://dx.doi.org/10.1111/j.1529-1006.2004.00016.x>

Kassin, S. M., Kukucka, J., Lawson, V. Z., & DeCarlo, J. (2014). Does video recording alter the behavior of police during interrogation? A mock crime-and-investigation study. *Law and Human Behavior*, *38*, 73-83. <http://dx.doi.org/10.1037/lhb0000047>

Kassin, S. M., Leo, R. A., Meissner, C. A., Richman, K. D., Colwell, L. H., Leach, A.-M., & La Fon, D. (2007). Police interviewing and interrogation: A self-report survey of police practices and beliefs. *Law and Human Behavior*, *31*, 381-400. <http://dx.doi.org/10.1007/s10979-006-9073-5>

Kleider, H. M., Pezdek, K., Goldinger, S. D., & Kirk, A. (2008). Schemadriven source misattribution errors: Remembering the expected from a witnessed event. *Applied Cognitive Psychology*, *22*, 1-20. <http://dx.doi.org/10.1002/acp.1361>

Lamb, M. E., Orbach, Y., Sternberg, K. J., Hershkowitz, I., & Horowitz, D. (2000). Accuracy of investigators' verbatim notes of their forensic interviews with alleged child abuse victims. *Law and Human Behavior*, *24*, 699-708. <http://dx.doi.org/10.1023/A:1005556404636>

Lassiter, G. D., Diamond, S. S., Schmidt, H. C., & Elek, J. K. (2007). Evaluating videotaped confessions: Expertise provides no defense against the camera-perspective effect. *Psychological Science*, *18*, 224-226. <http://dx.doi.org/10.1111/j.1467-9280.2007.01879.x>

Lassiter, G. D., Geers, A. L., Handley, I. M., Weiland, P. E., & Munhall, P. J. (2002). Videotaped interrogations and confessions: A simple change in camera perspective alters verdicts in simulated trials. *Journal of Applied Psychology, 87*, 867-874. <http://dx.doi.org/10.1037/0021-9010.87.5.867>

Lassiter, G. D., & Meissner, C. A. (Eds.). (2010). *Police interrogations and false confessions: Current research, practice, and policy recommendations*. Washington, DC: American Psychological Association. <http://dx.doi.org/10.1037/12085-000>

Leo, R. A. (1996). Inside the interrogation room. *Journal of Criminal Law and Criminology, 86*, 266-303.

MacLeod, C. M., Gopie, N., Hourihan, K. L., Neary, K. R., & Ozubko, J. D. (2010). The production effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 671-685. <http://dx.doi.org/10.1037/a0018785>

Morgan, C. A., III, Hazlett, G., Doran, A., Garrett, S., Hoyt, G., Thomas, P., & Southwick, S. M. (2004). Accuracy of eyewitness memory for persons encountered during exposure to highly intense stress. *International Journal of Law and Psychiatry, 27*, 265-279. <http://dx.doi.org/10.1016/j.ijlp.2004.03.004>

Neisser, U. (1981). John Dean's memory: A case study. *Cognition, 9*, 1-22. [http://dx.doi.org/10.1016/0010-0277\(81\)90011-1](http://dx.doi.org/10.1016/0010-0277(81)90011-1)

Schacter, D. L. (2001). *The seven sins of memory: How the mind forgets and remembers*. Boston, MA: Houghton Mifflin.

Schmidt, M. S. (2014, May 23). In policy change, justice dept. to require recording of interrogations. *The New York Times*, p. A14.

Snyder, C. J., Lassiter, G. D., Lindberg, M. J., & Pinegar, S. K. (2009). Videotaped interrogations and confessions: Does a dual-camera approach yield unbiased and accurate evaluations? *Behavioral Sciences & the Law, 27*, 451-466. <http://dx.doi.org/10.1002/bsl.875>

Sullivan, T. P. (2004). *Police experiences with recording custodial interrogations*. Chicago, IL: Northwestern University Law School, Center on Wrongful Convictions.

Sullivan, T. P. (2008). Recording federal custodial interviews. *The American Criminal Law Review*, 45, 1297-1345.

Sullivan, T. P., Vail, A. W., & Anderson, H. W. (2008). The case for recording police interrogation. *Litigation*, 34, 1-8.

The Justice Project. (2007). *Electronic recording of custodial interrogations: A policy review*. Washington, DC: Author.

*243 Thurlow, M. D. (2005). Lights, camera, action: Video cameras as tools of justice. *The John Marshall Journal of Computer & Information Law*, 23, 771-813.

Trainum, J. (2007, September 20). I took a false confession--so don't tell me it doesn't happen! *The California Majority Report*. Retrieved from <http://seeingtheforest.com/i-took-a-false-confession-so-dont-tell-me-it-doesnt-happen/>

Valentine, T., & Maras, K. (2011). The effect of cross-examination on the accuracy of adult eyewitness testimony. *Applied Cognitive Psychology*, 25, 554-561. <http://dx.doi.org/10.1002/acp.1768>

Appendix

Self-Report Questionnaire

1) In your opinion, is the suspect you just interrogated guilty or innocent?

(In your opinion, did the officer who interrogated you believe you were guilty or innocent?)

2) How confident are you in this opinion?

(How confident are you that this was his/her opinion?)

3) In your opinion, how credible were (the suspect's/your) denials?

4) How much did this suspect seem to know about the crime?

(In your opinion, how much did you know about the crime before you were questioned?)

5) In your opinion, how open and cooperative (was this suspect/were you) in (his/your) attitude and demeanor?

6) Did (this suspect/you) make any suspicious or self-incriminating remarks?

7) How hard did (you/*the officer*) try to get (this suspect/*you*) to confess?

[Phase 2 only: How much pressure did the police officer place on the suspect during the interrogation?]

8) How friendly, sympathetic, and understanding was (your behavior toward this suspect/*the officer toward you*)?

9) How confrontational (were you toward this suspect/*was the officer toward you*)?

10) How stressful do you think the interrogation was for this suspect?

(*How stressful was this interrogation for you?*)

11) Overall, how anxious (was this suspect/*were you*) during the interrogation?

12) Did (the suspect/*you*) make any admissions or confessions of guilt?

Received October 24, 2015

Revision received October 2, 2016

Accepted October 4, 2016

Footnotes

^{a1}

This article was published Online First December 12, 2016.

Saul M. Kassin, Department of Psychology, John Jay College of Criminal Justice; Jeff Kukucka, Department of Psychology, Towson University; Victoria Z. Lawson, Institute for State and Local Governance of the City University of New York; John DeCarlo, Department of Criminal Justice, University of New Haven.

This work was supported in full by a grant of the National Science Foundation, Division of Law and Social Sciences, to Saul M. Kassin (SES 1021422). We are also indebted to Kimberly Schantz and Stephanie Spiegel for their invaluable assistance in coding of the interrogations.

Correspondence concerning this article should be addressed to Saul M. Kassin, Department of Psychology, John Jay college of Criminal Justice, 524 West 59th Street, New York, NY 10019. E-mail: skassin@jjay.cuny.edu

41 LHUMB 230

End of Document

© 2017 Thomson Reuters. No claim to original U.S. Government Works.

97 B.U. L. Rev. 1157

Boston University Law Review

May, 2017

Symposium

The Fiftieth Anniversary of *Miranda v. Arizona*

Contribution

Christopher Slobogin ^{a1}

Copyright © 2017 by The Trustees of Boston University; Christopher Slobogin

MANIPULATION OF SUSPECTS AND UNRECORDED QUESTIONING: AFTER FIFTY YEARS OF MIRANDA JURISPRUDENCE, STILL TWO (OR MAYBE THREE) BURNING ISSUES

INTRODUCTION	1158
I. THREE GENERATIONS OF INTERROGATION TECHNIQUES	1160
II. THE LEGALITY OF SECOND-GENERATION TECHNIQUES	1164
A. <i>Coercion</i>	1164
1. Manipulative Techniques That Are Coercive	1168
2. Manipulative Techniques That Are Not Coercive	1171
3. The Rights Predicate and State Action	1173
4. Summary	1177
B. <i>Reliability</i>	1177
1. The Defendant's Burden of Production	1181
2. The Prosecution's Burden of Proof	1187
III. RECORDING	1188
A. <i>The Constitutional Arguments</i>	1190
B. <i>Extending the Recording Right Beyond the Stationhouse</i>	1192
IV. A WORD ON INTERROGATION OF SUSPECTED TERRORISTS	1194
CONCLUSION	1195

*Fifty years after Miranda, courts still do not have clear guidance on the types of techniques police may use during interrogation. While first-generation tactics (a.k.a. the third degree) are banned, second-generation tactics such as those found in the famous Reid Manual continue to be used by interrogators. The Supreme Court has sent only vague signals as to which of these second-generation techniques, if any, are impermissible, and has made no mention of newly developed third-generation tactics that are much less reliant on manipulation. This Article divides second-generation techniques into four categories: impersonation, rationalization, fabrication, and negotiation. After concluding, based on a review of field and laboratory research, that these techniques might well have superior “diagnosticity” to third-generation techniques--and thus that police might rationally want to continue using them-- *1158 it argues that the Court's Fifth Amendment and due process jurisprudence prohibits negotiation but permits impersonation, rationalization, and fabrication. At the same time, the Article recognizes that these techniques can produce false confessions; accordingly, it develops evidentiary principles for determining how courts might make use of expert testimony about factors that reduce the probative value of statements obtained during interrogation.*

To ensure the evidence necessary for this constitutional and evidentiary analysis, interrogations must be recorded. While a recording requirement has been endorsed by commentators from all points of the political spectrum, here too the Court has been silent. This Article summarizes why recording is required under the Due Process Clause, the Fifth Amendment, and

the Sixth Amendment, not just in the stationhouse but any time after custody. The Article ends with comments on how all of this should apply to interrogations of suspected terrorists. Together, these prescriptions give courts the concrete guidance the Supreme Court has failed to provide despite fifty years of case law.

INTRODUCTION

In the fifty years since *Miranda v. Arizona*¹ was decided, most of the big constitutional questions about interrogation have been resolved, at least to the Supreme Court's satisfaction. *Miranda* held that an individual's statements made during custodial interrogation are not admissible at a criminal trial unless the police first inform the individual about the right to silence and the right to counsel and unless the statements were made after a valid waiver of those rights.² The Court has since fleshed out the meaning of custody and interrogation, clarified the content of the warnings that must be given, decided numerous cases on the waiver issue, and limited the scope of the exclusion remedy.³ Certainly more case law from the courts, providing even greater nuance on these issues, is forthcoming. But the parameters have been set.

Yet the Supreme Court has still not explicitly addressed two issues that are crucial to sensible regulation of the interrogation process: the constitutionality of the various types of psychological techniques police often use during interrogation and the extent to which police must keep a record of the interrogation. Neglect of these two issues is particularly aggravating because, even prior to *Miranda*, they were arguably the most important outstanding questions relating to interrogation, other than *Miranda*'s core concern about whether notification of rights is constitutionally required. Well before *Miranda* was decided there was widespread consensus against the third degree and explicit threats designed to elicit confessions, meaning that for some time the *1159 key concern about the conduct of interrogations has been the propriety of using subtler psychological techniques.⁴ And any concrete rules developed on this score are likely to be toothless if accounts of the interrogation process are dependent entirely on ex post descriptions by the police and the suspect.⁵ Yet the Court, whether deliberating before or after *Miranda*, has offered only scattered dicta and vague hints about its views on the use of manipulative interrogation techniques.⁶ And despite the feasibility of recording interrogations for the past several decades, the Court has said nothing about the need for or even usefulness of recordings.⁷

This Symposium provides me with the opportunity to update articles I wrote addressing these two issues. The first article proposed what I called the “equivalency test,” which would permit police deceit about matters that, if true, would not be considered coercive, but did not address in detail how this test would interact with concerns about false confessions.⁸ The second article argued that the Due Process Clause, the Fifth Amendment, and the Confrontation Clause all require taping of a suspect's interrogation if the government later seeks to use those statements against the person in a criminal trial, but did not flesh out how this requirement would apply to statements made in the field.⁹

This Article expands upon these proposals by relating each to relatively recent developments. With respect to the use of manipulative interrogation techniques, it looks at three such developments: (1) the advent of “third-generation” *1160 interrogation practices that purport to avoid manipulation; (2) the explosion of research on techniques that may cause false confessions; and (3) conceptual advances regarding the admissibility of this research through expert testimony. Exploring these issues helps explicate how, under current doctrine, courts might more concretely evaluate the admissibility of self-incriminating information obtained through manipulative interrogation techniques.

With respect to the recording proposal, the update in this Article examines how the advent of police body cameras can transform regulation of pre-interrogation interrogation. Body cameras could allow courts to ensure that what transpires in the interrogation room is not a sanitized result of improper practices employed outside of it. This Article explains why, if body camera recording of suspect interviews is feasible, it is constitutionally required not only prior to criminal prosecution, but also prior to transport to the stationhouse.

Finally, this Article relates all of this to interrogations of individuals suspected of terrorism. Some have suggested that the usual restrictions on interrogation be relaxed in this context, given the public safety and national security concerns at stake.¹⁰ This Article argues that these government objectives can be met without changing interrogation rules.

I. THREE GENERATIONS OF INTERROGATION TECHNIQUES

Modern day interrogation relies on a host of techniques. While the physical abuse and prolonged confinement associated with the infamous “third degree”¹¹ is no longer officially sanctioned even by the police themselves,¹² a second generation of more psychologically manipulative strategies have taken their place. The most famous compendium of these techniques, purporting to have influenced hundreds of thousands of American police officers,¹³ is the manual published by John Reid and his coauthors.¹⁴ The practices recommended by Reid and his coauthors in cases of recalcitrant subjects almost all rely on some form of deception and can be categorized as follows: (1) “impersonation” (e.g., showing sympathy for the suspect, posing as a friend); (2) “rationalization” (e.g., suggesting that the confession will make the suspect feel better or appear honorable in the eyes of the community); (3) “evidence fabrication” (e.g., false statements that a codefendant has inculpated the suspect, that the suspect's fingerprints were found at the scene of the crime, and other means of insisting the suspect is guilty); and (4) “negotiation” (e.g., suggesting that, if the suspect confesses, more lenient punishment or release from detention is likely).¹⁵ These categories are fluid, and specific types of interrogation practices might sometimes fit in more than one of them, but, as will be demonstrated below, they are conceptually different from one another and come close to capturing the universe of manipulative interrogation conduct.

The most avid critics of police interrogation argue that none of these techniques should be permissible, on the ground that they are immoral, corrosive of the system, coercive, and liable to produce false confessions.¹⁶ Further, the argument is made that none of these techniques is necessary, given equally or more effective “third generation” alternatives.¹⁷ The two most prominent alternatives are the “PEACE” method (for Preparation and Planning; Engage and Explain; Account; Closure; and Evaluation),¹⁸ and the approach developed by the High-Value Interrogation Group Research Unit (“HIG”) established by President Obama in 2010 as part of the counter-terrorism effort.¹⁹ PEACE, purportedly in use in the United Kingdom, New Zealand, Norway and parts of Canada,²⁰ eschews all of the techniques described above, and relies on confronting the suspect with valid evidence and discrepancies in his or her story.²¹ The HIG technique is very similar to PEACE but relies more explicitly on practices that produce “cognitive load,” the idea that liars have a much harder time than truth-tellers at keeping their story straight.²² Thus, the goal with HIG is to develop such rapport with suspects that they feel comfortable talking, with the result that they eventually trip over their own lies.²³

Both of these approaches, and in particular the HIG approach, rely heavily on convincing the suspect that the interrogators are trustworthy.²⁴ HIG interlocutors are also taught to rely on what some might call “tricks,” such as asking suspects to describe the relevant events in reverse order, maintain eye contact with the interrogator, perform several tasks at once, and answer unexpected questions, all designed to increase cognitive load.²⁵ A variant of the HIG approach, known as the Scharff technique after Hanna Joachim Scharff, a German World War II interrogator, also relies heavily on the “illusion of knowing it all.”²⁶ Thus, even some third-generation approaches rely on at least a junior form of deception, and may be suspect to those who believe interrogation should not depart from Sunday School norms.²⁷

More importantly, the jury is still out as to how effective these third-generation tactics are. It has been asserted that confession rates achieved using the PEACE technique are at least as high as those in the United States.²⁸ But that *1163 claim can be disputed on a number of grounds.²⁹ A stronger claim, based on research to date, is that third-generation interrogation techniques produce a greater amount of information than more confrontational practices that might be met with denials and clam-ups.³⁰ That is a decided advantage if, as the cognitive overload theory predicts, more information means a greater chance of exposing contradiction.³¹ A final claim that is likely true is that, given their relatively unaggressive posture, third-generation techniques are less likely to cause false confessions.³²

Nonetheless, yet to be proffered is solid evidence that, compared to second-generation techniques, third-generation techniques have superior “diagnosticity”—that is, a similar or higher true confession rate combined with a lower false confession rate.³³ Second-generation techniques may be significantly better than third-generation techniques at producing true confessions,³⁴ and if any false confessions they generate can be exposed prior to *1164 use in adjudication (in ways explained later in this Article³⁵), their ultimate diagnosticity may be superior. If so, American police are justified in refusing to abandon second-generation techniques on effectiveness grounds. Then the key question is whether these techniques are legal.

II. THE LEGALITY OF SECOND-GENERATION TECHNIQUES

Until the superiority of third-generation interrogation techniques becomes clear, the primary issue surrounding interrogation will be whether there are any legal impediments to second-generation practices. Below, that inquiry is divided into an exploration of the Fifth Amendment's prohibition on coercion and an examination of the relevance of confession reliability under both the Constitution and evidence rules. The goal of this discussion is conservative: it is to make sense of current doctrine, not change it—to help courts think through the implications of the Court's cases, not propose a new framework. The discussion of coercion is an amplification of my previous work on that topic, while the discussion of reliability brings together an analysis of the Supreme Court's decision in *Colorado v. Connelly*,³⁶ a brief description of the new research findings from social psychologists about possible causes of false confessions, and some thoughts on how those findings can best be used by the courts.

A. Coercion

The Due Process Clause has long protected against “involuntary” confessions.³⁷ *Miranda* broadened the definition of coercion by moving the *1165 inquiry from the Due Process Clause to the Fifth Amendment's prohibition on compelled self-incrimination.³⁸ Relying on that language, *Miranda* reoriented the concept of coercion in the interrogation context, holding that compulsion occurs during any custodial questioning that is not preceded by the famous warnings about the rights to silence and counsel.³⁹ At the same time, *Miranda* held that if police provide the four-part litany and the suspect says he understands it and decides to speak anyway, then, barring coercive conduct by the police, any statements subsequently made by the defendant should be admissible under the Constitution.⁴⁰ One can argue whether the Fifth Amendment in fact requires this regime,⁴¹ but this Article will take it as a given.

The Court's interrogation case law has always been vague about the precise meaning of coercion, understandably so given how that concept has perplexed moral philosophers.⁴² Even so, some concrete guidance can be extracted from the Court's jurisprudence. The physical abuse or prolonged (multi-day) detention of suspects associated with first-generation practices is clearly coercive under the Constitution.⁴³ Similarly, when police threaten a suspect's *1166 loved one with

physical force or other serious harm, then any ensuing confession is involuntary.⁴⁴ In *Miranda's* terms, these situations impose too high a cost on remaining silent, effectively nullifying the right.⁴⁵

The Court has not been as clear about the propriety of second-generation tactics, however. In several pre-*Miranda* cases, where the suspect was subjected to custodial interrogation and yet was not told about the right to remain silent, the Supreme Court was leery of confessions obtained through deception.⁴⁶ In contrast, in post-*Miranda* cases where the *Miranda* litany was given, the Court has appeared to be more willing to countenance manipulation by the police.⁴⁷ Certainly that has been true in some lower courts, which have permitted all four types of second-generation tactics described above: impersonation,⁴⁸ *1167 rationalization,⁴⁹ fabrication,⁵⁰ and negotiation.⁵¹ The fact remains that the Supreme Court itself has been coy about these matters.⁵²

Nonetheless, more specific guidance, consistent with the Court's case law such as it is, is possible. In *Lying and Confessing*, I argued that while first-generation and negotiation techniques are impermissibly coercive, impersonation, rationalization, and evidence fabrication are not.⁵³ Following the lead of the Court, I did not derive this conclusion from philosophical musings. Rather, I based it on the simple precept that, once the warnings are given and acknowledged as understood, police deception during interrogation amounts to Fifth Amendment coercion when, but only when, the deceptive statements would be coercive if true, a principle I called the "equivalency rule."⁵⁴ A police statement that, if true, is not unconstitutionally coercive, does not become coercive simply because it is in fact false. Conversely, of course, all coercive tactics, whether deceptive or not, should be banned. The implications of this equivalency test, and how it ties in with Supreme Court precedent governing the interrogation process, are explored below.

*1168 1. Manipulative Techniques That Are Coercive

The most obviously coercive deceptive practices under the equivalence test are statements that a suspect's postwarning silence will be used against him or that contact with counsel will be prevented, because even if the police mean what they say these declarations are a direct violation of *Miranda*.⁵⁵ Also clear is that false threats to impose a legal penalty if a confession is not forthcoming are coercive, as these threats would be coercive if true. The Supreme Court's Fifth Amendment jurisprudence has long prohibited imposition of legal sanctions for refusing to make self-incriminating statements.⁵⁶

Of course, police rarely are so blatant. More likely are statements indicating either that "things will get worse" if silence or counsel rights are asserted,⁵⁷ or the converse, that lenient legal treatment is likely if the suspect confesses.⁵⁸ Often these descriptions of the suspect's legal plight might turn out to be true.⁵⁹ Nonetheless, whether true or false, such negotiation techniques tell the suspect that remaining silent will, in effect, result in a criminal penalty. Under the Fifth Amendment they should be considered coercive. Even the Reid Manual emphasizes that suspects should not be told that they face certain punishment if they do not confess or promised leniency if they do.⁶⁰

*1169 Again, however, sophisticated police are not likely to resort to such tactics. Instead, one technique that the Reid Manual recommends is to intimate that the suspect will have some type of defense if he or she confesses.⁶¹ Or police might suggest that the suspect will get to go home if incriminating information is provided, or engage in extremely long interrogations that imply the same thing.⁶² The inquiry here becomes more difficult, but should ultimately depend on the extent to which police condition better legal treatment on a confession. For instance, questions such as "Have you done this many times before or was this just the first time?" or "Was this whole thing your idea or did you get talked into it?"⁶³ are much closer to rationalization than negotiation techniques; they give the suspect a reason to feel less guilty about the offense, but do not suggest that real legal consequences will flow from a confession. In contrast, both direct

and indirect indications that a confession will mean more lenient treatment by the court (e.g., “You are not to blame, but you have to tell my why”) or by the police (e.g., “We can make this short or long”) should lead to exclusion.⁶⁴ In such cases the police are telling the suspect that a confession is the only way to avoid significant criminal liability or physical detention.

*1170 This admonition is also consistent with the Supreme Court's cases. In the late nineteenth century case of *Bram v. United States*,⁶⁵ the Court held invalid under the Fifth Amendment any confessions “extracted by any sort of threats or violence, ... obtained by any direct or implied promises, however slight, [or] by the exertion of any improper influence.”⁶⁶ Of course, the suspect in *Bram* did not have the benefit of *Miranda* warnings. However, over seventy years later, a post-*Miranda* case, *Brady v. United States*,⁶⁷ explained that in *Bram* “even a mild promise of leniency was deemed sufficient to bar the confession, not because the promise was an illegal act as such, but because defendants at such times [alone and unrepresented by counsel] are too sensitive to inducement and the possible impact on them too great to ignore and too difficult to assess.”⁶⁸ Seven years after *Brady*, in *Hutto v. Ross*,⁶⁹ the Court implied that a counsel-less plea deal conditioned on a confession violated that test.⁷⁰ While the subsequent decision of *Arizona v. Fulminante*⁷¹ stated that “under current precedent [*Bram*] does not state the standard for determining the voluntariness of a confession,”⁷² that case dealt solely with due process analysis outside of the custodial context. In contrast, direct or implied promises during custodial interrogation that condition silence on a legal penalty, whether true or false and whether pre- or postwarning, directly violate the Fifth Amendment's commands. Many lower courts, although certainly not all, are in accord.⁷³

One defense of negotiation techniques--at least those that focus on promises of legal leniency--is that they are very similar to the process of plea bargaining,⁷⁴ which the Supreme Court has enthusiastically sanctioned.⁷⁵ It is *1171 true that the Supreme Court has held that “[d]efendants advised by competent counsel and protected by other procedural safeguards are presumptively capable of intelligent choice in response to prosecutorial persuasion, and unlikely to be driven to false self-condemnation.”⁷⁶ But, as this language makes clear, the legality of plea bargaining is dependent on the participation of counsel, both the defense attorney and the prosecutor, as well as the supervision of the judge at the plea colloquy. More importantly, in terms of Fifth Amendment compulsion concerns, “pre-plea bargaining” is unconstitutional because of the implicit or explicit message that if counsel is consulted, the deal is off the table; that message directly undercuts both the right to silence and the right to counsel.⁷⁷

2. Manipulative Techniques That Are Not Coercive

While the equivalency principle bars any deception that directly undercuts the warnings, it permits many other types of manipulation, including impersonation, rationalization, and fabrication. These techniques often or always involve deceptive statements by the police. But, in line with the few hints we have from Court's cases,⁷⁸ they are not coercive because the same statements would be uncoercive if true.

Take impersonation, or what Welsh White has called the “pretended friend” technique.⁷⁹ Officers expressing sympathy for the suspect's plight or pretending to be the suspect's new best friend can be highly deceptive.⁸⁰ But they are not acting coercively, or at least no more coercively than a friend acts. Whether or not the interrogator is in fact a friend or colleague in crime, the pressure to talk in this situation is virtually nonexistent. These scenarios merely encourage the suspect to, as Bill Stuntz put it, “forget” about the existence of the right encapsulated in the warnings.⁸¹ The friendly cop might also be joined by a *1172 tougher one, as in the infamous Mutt and Jeff routine.⁸² However, so long as Mutt does not engage in the third degree or negotiation techniques and Jeff does not offer protection from those techniques or offer leniency, Fifth Amendment coercion has not occurred. To the extent they rely on “fake” trust, this conclusion insulates the PEACE and HIG approaches as well.

The equivalency test would also permit rationalization tactics. For instance, police might suggest that there are psychological benefits to confessing, such as alleviating feelings of guilt, showing concern for the victim's family, assuring forgiveness from God, or achieving respect in the community.⁸³ These sentiments would not be considered coercive if the police sincerely voiced them. That they are often voiced pretextually should not change the analysis. The important caveat is that such techniques cannot merge into negotiation tactics promising legal relief, because then they become coercive in the Fifth Amendment sense.⁸⁴

More controversially, the equivalency test sanctions evidence fabrication ploys. Confronting a suspect with actual forensic evidence discovered at the crime scene, actual eyewitness accounts, or actual documentary evidence obviously produces pressure to confess (in the case of guilty people) or explain (in the case of innocent ones). But if such tactics were considered unconstitutionally coercive, even the PEACE approach would have to be outlawed, and confessions triggered by evidence that later turns out to be wrong (e.g., an incorrect eyewitness identification) would have to be thrown out as well. If instead the evidence is made-up, the pressure to talk is, at worst, usually no more intense and perhaps even reduced, since the suspect, whether guilty or innocent, can often smell out the ruse. For the same reason, police “bluffing” about how certain they are that the suspect is guilty should not be considered coercive.⁸⁵ Even if the police go to the trouble of fabricating evidence that can *1173 be shown to the suspect, the pressure to talk is no greater than in cases where the evidence actually exists.⁸⁶ Courts might bar such tactics on the ground that the fabricated evidence will confuse a jury if it must be presented at trial to explain how the confession was obtained.⁸⁷ But that decision would be based on concerns about prejudicing defendants, not coercing them.⁸⁸

3. The Rights Predicate and State Action

Under the equivalency test, lies about the rights encased in the warnings constitute per se coercion under the Fifth Amendment.⁸⁹ Less clear is the result if the police neither lie nor mislead about the *Miranda* rights, but rather underplay them, or allow misimpressions on the part of the suspect to continue uncorrected. The discerning reader may have noted that, in my earlier description of *Miranda*'s holding,⁹⁰ I said that *Miranda* prevents the police from *telling* suspects they do not have a right to silence and counsel or that their statements will not be used against them, and that *Miranda* requires that police make sure the suspect *says* he understands these rights. I described the holding this way because, since *Miranda*, the Supreme Court has held admissible a number of confessions obtained after the police accurately give the warnings to a suspect who indicates he understands them and then gives incriminating statements, all while apparently confused about what the warnings really mean. *1174 For instance, in *North Carolina v. Butler*⁹¹ and *Connecticut v. Barrett*,⁹² the defendant appeared to believe that statements not reduced to writing and signed are inadmissible; in *Berghuis v. Thompkins*⁹³ the defendant might have believed that merely remaining silent meant he had asserted his right to silence;⁹⁴ in *Davis v. United States*⁹⁵ the defendant probably thought that stating “Maybe I should talk to a lawyer” was an assertion of the right to counsel;⁹⁶ and in *Colorado v. Spring*⁹⁷ the defendant may have believed that he could not refuse to answer a question about a murder after talking volubly about a firearms violation.⁹⁸ In all of these cases, the Court found no violation of *Miranda*.

Assuming that during the interrogation the suspects had the beliefs just ascribed to them, were their statements coerced under the Fifth Amendment? The dissenters in some of these cases and many commentators believe so, on the ground that any confusion about the *Miranda* rights means that subsequent statements are compelled.⁹⁹ But this commentary has given insufficient *1175 consideration to the implications of *Colorado v. Connelly*, which (correctly) held that neither the Fifth Amendment nor the Due Process Clause is violated unless the police engage in “overreaching” that leads to the statement.¹⁰⁰ All of the cases under consideration involved inaction, not action; the pressure to talk, if there was any

(unlikely in *Butler* and *Barrett*, possible in the other three), came not from the police but from the defendant's confusion about rights that had been read to him and that he said he understood. The police did not “overreach,” they merely took advantage of a suspect's befuddlement.¹⁰¹ The one possible exception is *Spring*, where the police affirmatively lied about the full scope of the interrogation, probably in an effort to surprise Spring with their question about the murder midway through the interview.¹⁰² But that lie was not about the rights to silence or to counsel. While its timing might have taken advantage of Spring's erroneous belief that once he started talking he had to keep going, that belief was Spring's “fault,” not the fault of the police.

A second non-doctrinal, but perhaps equally important, reason to accept the results in these cases is the heavily documented fact that a large proportion of suspects have trouble understanding the warnings.¹⁰³ Imposition of a duty to clarify would place substantial burdens on the police. Further, when a suspect claims the rights should have been clarified, courts must also determine whether any claimed misunderstanding was real. It was assumed above that the defendants in *Butler*, *Barrett*, *Berghuis*, *Davis*, and *Spring* thought either that nothing they said could be used against them (in *Butler* and *Barrett*) or that they did not have a right to silence or counsel (in *Berghuis*, *Davis*, and perhaps *Spring*). But in many such cases there will be evidence to the contrary.¹⁰⁴ If so, *1176 the ensuing suppression hearing would require ascertaining not only whether the police should have been on notice that the defendant might need clarification, but also whether the defendant actually needed the clarification; the temptation to malingering confusion is very high in such situations.

In short, despite their underhanded treatment of the *Miranda* warnings, these cases make sense both as a matter of doctrine and as a practical matter. Closer cases occur when police immediately proceed to questioning after giving the warnings and eliciting an indication of understanding without asking the suspect whether he or she wants to invoke the rights, or when they downplay the rights as mere bureaucratic boilerplate.¹⁰⁵ In these situations more is involved than a failure to clarify, and a court might find that police affirmatively led the suspect to believe he was supposed to talk. But even in these situations the conclusion that silence is not permissible is the suspect's; the police are not stating or implying there will be a legal penalty for remaining silent.

More easily distinguishable from the Court's cases are those situations where psychological characteristics of the suspect make the warnings irrelevant. This situation is most likely to arise with very young children (below thirteen) or individuals with intellectual disability. Because of their susceptibility to authority figures and their difficulty understanding abstract concepts, they are very likely to believe that they should talk to police regardless of how carefully the *Miranda* warnings are delivered.¹⁰⁶ In other words, the mere act of questioning these sorts of people “compels” them to talk despite being told about the right to silence. Further, in contrast to intellectually intact adults, whose actual understanding of the rights can be difficult to discern, the cognitive *1177 deficits of young children and people with intellectual disability are consistent and relatively reliably ascertained by appropriately trained individuals.¹⁰⁷

4. Summary

One way of deciding whether manipulative interrogation techniques are permissible is to analyze whether they would be coercive if police were in fact acting in good faith. Under this equivalency test, third degree interrogation and negotiation tactics involving threats or promises about a suspect's legal situation should be banned under the Fifth Amendment. Impersonation, rationalization, and fabrication should not be. The legal effect of a failure to clear up confusion about rights once they have been recited and the suspect claims to understand them is less clear but can be justified on the ground that the failure does not amount to state action.

Of course, most interrogations, at least those that are contested, usually involve a mishmash of techniques, not just a single tactic, often over a several-hour period.¹⁰⁸ Figuring out which technique, if any, “caused” an incriminating

statement is virtually impossible. Probably the best approach is to presume that any interrogation in which negotiation tactics or first-generation techniques are used is coercive for Fifth Amendment purposes. The prosecution bears the burden of disproving coercion by a preponderance of the evidence.¹⁰⁹ Unless the prosecution can proffer solid evidence that coercive tactics, once shown to have occurred, did not influence the suspect, the confession should be excluded on Fifth Amendment grounds.

B. Reliability

The conclusion that confessions that are coerced should be excluded follows even if the prosecution can convincingly show the confession is reliable. The Fifth Amendment prohibits compelled testimony, not just compelled, unreliable testimony.¹¹⁰ The converse of that statement is not true, however. As its language makes clear, and as the Supreme Court has confirmed, the Fifth *1178 Amendment has nothing to say about false confessions unrelated to compulsion.¹¹¹

Nor, as a practical matter, does the Due Process Clause. Some Supreme Court cases prior to *Miranda* did refer to the potential role of the Clause in excluding false confessions independently of whether they were coerced.¹¹² But in *Colorado v. Connelly* the Court not only emphasized the state action requirement, it also rejected this earlier view of the process due during interrogation.¹¹³ After noting that the lower court in that case had found the confession resulted from the defendant's mental conflicts rather than from police interrogation,¹¹⁴ the majority in *Connelly* declared that “[a] statement rendered by one in the condition of respondent might be proved to be quite unreliable, but this is a matter to be governed by the evidentiary laws of the forum, ... not by the Due Process Clause of the Fourteenth Amendment.”¹¹⁵ Thus, the Court dismissed the lower court's holding that the Clause requires “inquiries ... divorced from any coercion brought to bear on the defendant by the State.”¹¹⁶

This language also strongly suggests that, whatever may have been true decades ago,¹¹⁷ due process analysis in interrogation cases is no longer concerned with “offensiveness” other than that associated with coercion. Pre-*Miranda* cases, where warnings were not required, excluded confessions on the ground that the techniques used were “revolting to the sense of justice,”¹¹⁸ *1179 “tyrannical,”¹¹⁹ or “shock[ing to] the conscience.”¹²⁰ But this type of language--which is extremely amorphous,¹²¹ and partly for that reason has generally been avoided in other contexts¹²²--rarely finds its way into post-*Miranda* cases.¹²³ It is possible that particularly egregious police interrogation conduct offends the Due Process Clause even if it is not coercive.¹²⁴ But such cases are virtually an extinct breed after the *Connelly* decision and the Court's repeated description of the due process test as one focused on interrogation practices “calculated to break the suspect's will.”¹²⁵ In short, for all practical purposes, the protection afforded by the Due Process Clause and the protection guaranteed by the Fifth Amendment after the warnings are given and putatively understood are co-extensive.¹²⁶

*1180 If so, proof that a confession is false does not change the constitutional analysis in interrogation cases.¹²⁷ That does not mean that defendants cannot obtain exclusion of false confessions, of course. As *Connelly* indicates, local evidentiary rules are another basis for exclusion.¹²⁸ Every jurisdiction provides that evidence lacking in probative value or whose probative value is outweighed by its prejudicial impact is inadmissible.¹²⁹ Certainly, false confessions, which are almost always completely inaccurate at the same time they are extremely influential,¹³⁰ fall in that category.¹³¹ Commentators have also noted that false confessions might be excluded under the rule that a witness may only testify based on personal observation,¹³² under variations of the *corpus delicti* rule,¹³³ *1181 or pursuant to a special rule

of evidence focused solely on confessions.¹³⁴ Procedurally, these claims would be raised via a motion *in limine*, akin to a suppression hearing.¹³⁵

While an evidentiary hook is necessary,¹³⁶ the more important question is the quantity and quality of evidence defendants must present to meet the burden of production. On the assumption that the Due Process Clause protects against false confessions even in the absence of coercion, Eve Primus proposes that the defense should be required to show that the police used techniques they knew or should have known were likely to cause a false confession, at which point the state must show by a preponderance that the confession was reliable.¹³⁷ Because, as just explained, proof of state action is not necessary in the evidentiary setting, neither is proof of police mens rea of the type required by Primus.¹³⁸ But otherwise this division of responsibility between the defense and the prosecution is a sensible proposal, which the following discussion fleshes out in more detail.

1. The Defendant's Burden of Production

In meeting its burden of production under evidence law, the defense must provide plausible evidence that the defendant's confession is false. In the absence of significant physical or eyewitness proof of innocence (which is likely if the case continues forward), this showing can be difficult. Fortunately, because of the huge increase in relevant social science research, defense attorneys have been able to rely on more than conjecture for this purpose. In the past two decades, social scientists have conducted studies purporting to find a large number of “risk factors” for police-induced false confessions, including: (1) bargaining techniques involving legal consequences; (2) minimization of guilt that falls short of promising legal leniency; (3) interrogations lasting over four to six hours; (4) sleep deprivation; (5) false evidence ploys, especially when combined with lengthy interrogations; and (6) “bluffing” to the suspect that untested forensic evidence exists, which can induce a belief that exoneration will *1182 occur despite a confession.¹³⁹ Then there are a number of other risk factors related to the suspect rather than to specific conduct by the police: (7) the belief that the criminal justice system is fair and thus exoneration is forthcoming; (8) immaturity; (9) intellectual disability; and (10) mental illness, including antisocial personality disorder, psychosis, and depression.¹⁴⁰ The basic research underlying these last four findings, particularly the latter three, documents well-known psychological mechanisms, including difficulty in delaying gratification, susceptibility to suggestion, and vulnerabilities in memory.¹⁴¹

Some of these risk factors (most obviously, (1) and, depending on the circumstances, (3), (8), and (9) as well) overlap with techniques or dispositions that would require a finding of coercion under the Fifth Amendment.¹⁴² But many would not. If an interrogation was not coercive under the Fifth Amendment and the defense instead wants to exclude the confession on the ground it is false, how might this evidence be presented? Answering this question requires resort to the rules of evidence regarding expert testimony, which today center on *Daubert v. Merrell Dow Pharmaceutical Co.*¹⁴³ Although these rules are often relaxed in pretrial hearings,¹⁴⁴ they need to be discussed in this setting because they clearly have affected appellate court decisions about the admissibility of expert evidence pertaining to false confessions, which often sustain exclusion of this type of testimony whether presented prior to or during trial.¹⁴⁵

Primus notes that such testimony can be “generalized” or “particularized” but does not discuss the evidentiary implications of this distinction.¹⁴⁶ In a recent article entitled *Group to Individual (G2i) Inference in Scientific Expert Testimony*, David Faigman, John Monahan, and I argued that *Daubert* analysis should differ depending upon whether the testimony is about general *1183 phenomenon (i.e., “framework” evidence) or individual characteristics (i.e., “diagnostic evidence”).¹⁴⁷ While the difference in analysis between framework and diagnostic evidence can vary along five axes (relevance, qualifications, internal validity, helpfulness, and prejudicial impact),¹⁴⁸ the relevance and validity components are most pertinent here.

In the confession context, framework evidence would report general research findings of the type described above and let the factfinder draw conclusions about whether it applies to the case at hand. One advantage of framework evidence is that because it is generally applicable, reaching beyond the facts of a particular case, it can be presented in briefs, the same way generally applicable legal principles are proffered;¹⁴⁹ in this way, litigants who cannot afford an expert witness might still be able to take advantage of false confession research. A possible disadvantage of framework evidence is that its general nature can sometimes undermine its relevance to a particular case, an issue that we referred to as “empirical fit”¹⁵⁰ (borrowing from *Daubert's* use of the latter word¹⁵¹) and that social scientists call external validity or generalizability.¹⁵² Whatever the label, the concept refers to the extent to which research findings apply to groups or individuals that were not the subject of study.¹⁵³

For much of the research on false confessions, empirical fit may be unclear because it relies on “interrogations” of college students and similar populations who are accused of minor infractions that at most will lead to some type of academic penalty, and who are not given *Miranda*-style warnings.¹⁵⁴ Not surprisingly, the generalizability of these findings to warned criminal defendants *1184 charged with serious crimes and potentially subject to imprisonment has been called into question.¹⁵⁵ Further, because they take place in the “lab” and must abide by research ethics standards, the studies have a hard time replicating the incentives of real criminal defendants. In the popular “computer-crash” paradigm, subjects are falsely informed they have caused a computer to crash by pressing a button they were told not to press.¹⁵⁶ While a large number of these innocent subjects “confess,” they could easily be doing so because they believe they are in fact guilty, given the vagaries of typing (as indicated by the finding in the most famous such study that the researchers obtained confessions from 65% to 100% of those in the “fast-paced” condition but from only 35% to 89% of those under the “slow-paced” condition).¹⁵⁷ The better constructed “cheating” paradigm, where researchers obtained confessions from students who in fact did not cheat and presumably knew they did not, largely avoids that problem, and research using it tends to corroborate that minimization, false evidence, and bluffing techniques increase false confessions.¹⁵⁸ But the findings in these studies that even those not subject to any manipulative questioning sometimes falsely confess (at a rate ranging from 6%¹⁵⁹ to 26.7%¹⁶⁰) indicates that students *1185 may not consider the consequences of cheating during an experiment that significant.

Research on interrogations involving actual criminal defendants has greater empirical fit or relevance. But it can suffer from suspect internal validity, because the “ground truth” of whether a confession is in fact false can only be known in a small subset of cases.¹⁶¹ Some researchers have avoided this problem by focusing on proven cases of wrongful conviction in which confessions were obtained, and they have produced work that suggests a correlation between false confessions and techniques such as negotiation, evidence ploys, and lengthy interrogation.¹⁶² However, because these latter techniques occur in a large number of interrogations, most of which produce true confessions or at least confessions not known to be false, this research is still ambiguous about whether the techniques studied are likely to lead people to confess to crimes they did not commit.¹⁶³ The fact that ninety percent of false confessions came after interrogations lasting more than three hours, reported by Brandon Garrett,¹⁶⁴ does not mean that most such interrogations, or even a sizeable minority of them, produce false confessions.

Despite these external and internal validity problems, where the laboratory and field research is convergent--say, with respect to the impact of minimization techniques or the combined impact of the false evidence ploy and lengthy interrogation¹⁶⁵--this type of evidence should probably be admissible, *1186 especially since it will usually also be very helpful (another evidentiary factor), in the sense that it challenges preconceptions about the likelihood that an innocent person would confess.¹⁶⁶ However, the court must still decide whether, in light of the expert evidence, the defendant's burden of production is met. For instance, one laboratory study suggests that some techniques increase the risk of a false

confession three-fold, from six percent to eighteen percent.¹⁶⁷ But even ignoring external validity concerns, if in the real world the base rate for false confessions in unmanipulated interrogations is infinitesimal, a three-fold increase due to police manipulation might not be considered significant.¹⁶⁸

This is where diagnostic expert testimony about a particular defendant's mental condition and reaction to police conduct at the time of the interrogation could play a role. This testimony too must satisfy evidentiary requirements. In contrast to framework testimony about confessions, diagnostic testimony is most likely to be vulnerable on validity rather than relevance grounds. In *Group to Individual Inference*, we pointed out that the ideal method of determining the validity of diagnostic testimony is through feedback loops that provide data about the accuracy of a particular expert's conclusions.¹⁶⁹ Unfortunately, such feedback is unlikely in the false confession context. Alternatively, we argued, the validity of diagnostic evidence can be improved through ensuring the expert's assessment is based on an empirically derived, structured evaluation process.¹⁷⁰ Social scientists have developed psychometrically sound instruments that can help measure a defendant's understanding of *Miranda* and his or her suggestibility.¹⁷¹ Although the extent to which such instruments assess confession reliability is not clear, evaluations based on these or similar protocols can suggest that, due to youth, mental disorder, or other personality traits, the defendant exhibits significant cognitive impairment, suggestibility, or impulsivity. Those types of results, combined with the relevant framework evidence, could easily lead a court to decide that the probative value of a given confession is so low it is outweighed by the potential the confession will blind the jury to evidence of innocence. If the court so finds, the confession should be declared inadmissible unless the prosecution can burnish its probative value.

2. The Prosecution's Burden of Proof

Assuming the prosecution does not have overwhelming evidence independent of the confession (in which case the confession is unnecessary), how can it rebut a finding that a confession is presumptively false? Richard Leo and Richard Ofshe have suggested three factors must be considered: (1) whether the confession contains nonpublic information that can be independently verified, would be known only to the true perpetrator or an accomplice, and cannot likely be guessed by chance; (2) whether the confession led the police to new evidence about the crime; and (3) whether the suspect's postadmission narrative fits the crime facts and other objective evidence.¹⁷² Proof of any of these factors will go a long way toward showing the confession is reliable, with the important caveat that the prosecution must also show that interrogators did not feed the suspect the relevant information or simply fraudulently assert that he or she knew it.¹⁷³

If a suspect confesses without any detail, or the crime is mundane enough that there are no special facts, the prosecution's burden on this score is more difficult.¹⁷⁴ But if Leo and Ofshe's proposal becomes the rule, police should be ***1188** required to act accordingly. A confessor who remains otherwise mute should be cajoled into providing detail that can be corroborated with nonpublic information, used in discovering new evidence, or compared to known facts.¹⁷⁵ Police should make sure they do not contaminate a confession by providing the suspect with information only the perpetrator would know or, if their confrontation tactics make that impossible, at least withhold one such piece of information until a confession is forthcoming so as to provide a double-check.¹⁷⁶ These procedures are simple enough that, without a very good explanation for why they were not followed, the prosecution should not be able to meet its burden on the reliability issue as a matter of law.

III. RECORDING

None of the foregoing determinations about coercion and reliability can be made with any confidence without a recording of the interrogation, preferably on video. Unencumbered by real-time depiction of their conduct, police may be very reticent about admitting to engaging in negotiation and other manipulative techniques, and anything suspects say on that

score will look self-serving and thus often lack credibility.¹⁷⁷ Even if both parties agree about what happened *1189 and give accurate accounts to the best of their ability, subtleties about impersonation, rationalization, evidence fabrication, and negotiation will be missed; confession contamination may also be hard to discern without a record.¹⁷⁸ In civil cases, face-to-face questioning of one party by the opposing party, conducted with the goal of producing evidence for trial, virtually always takes place at a deposition, and a deposition that is not recorded in some fashion is always inadmissible evidence.¹⁷⁹ This is in stark contrast to interrogations conducted in the criminal justice system. As I stated in *Toward Taping*, “it is stunning that we do not require verbatim transcripts of criminal interrogations, where the stakes are so much higher, access to information about psychological pressures so much more important, and legal representation (of either party) so much less likely.”¹⁸⁰

Motivated as much by a desire to deter defendants from making up stories as by the objective of providing the courts with evidence, many police departments are moving toward recording interrogations.¹⁸¹ But many have not done so, and the effort is often half-hearted; interrogations at the stationhouse may not be recorded in full, and any softening up of the suspect prior to arrival at the *1190 stationhouse is virtually never subject to recording.¹⁸² A constitutional argument is needed to ensure recording takes place. In *Toward Taping*, I provided three.¹⁸³ After canvassing those arguments, which were aimed at stationhouse recording, I extend the analysis to pre-interrogation interrogation.

A. *The Constitutional Arguments*

The first constitutional argument that interrogations must be recorded is based on the Due Process Clause and straightforwardly asserts that procedural fairness requires a recording.¹⁸⁴ This argument must clear the hurdle created by the Supreme Court's decisions in *California v. Trombetta*¹⁸⁵ and *Arizona v. Youngblood*,¹⁸⁶ which held that failure to preserve forensic evidence after it has been tested does not violate due process unless the defendant is denied access to the test results or bad faith is otherwise proven.¹⁸⁷ Relying on these cases, a number of lower courts have held that a failure to record an interrogation is not a violation of due process because it is not designed to hide exculpatory information and the defendant can reconstruct the interrogation through testimony from the suspect and the police.¹⁸⁸ But these courts misconceive the *1191 problem. Failing to tape is much worse than destroying physical evidence that has been tested because, in the interrogation setting, until a court looks at the interrogation transcript, the evidence has yet to be “tested;” just as is true of untested forensic evidence, the defendant's only evidence in this instance is his or her say so. Further, as already noted, neither defendants nor courts can accurately reconstruct the interrogation based solely on the testimony of the interrogators and the suspect.

The second argument is an originalist one based on the Fifth Amendment. At the time the Amendment was drafted and well afterward, all interrogations were conducted by a judge in open court.¹⁸⁹ Neither police departments nor their interrogation rooms existed.¹⁹⁰ Assuming we are not going to move back to the colonial model, the closest modern equivalent to such questioning is a recording that allows judges to witness the interrogation as it happened.¹⁹¹

The third argument in support of recording is grounded in the Sixth Amendment decision in *United States v. Wade*,¹⁹² which held that defendants subjected to lineups are entitled to counsel or “substitute counsel.”¹⁹³ Although *Wade* is usually described as a case about the right to counsel, a more accurate reading of the case is that it rests on a separate Sixth Amendment right, the right of confrontation. Worried about the “vagaries” of eyewitness identifications,¹⁹⁴ the Court emphasized that without some third-party mechanism for recounting how the lineup occurs the accused is “deprived of that right of cross-examination which is an essential safeguard to his right to confront the witnesses against him.”¹⁹⁵

Most lower courts have interpreted *Wade* to require, at the least, a visual *1192 depiction of the lineup.¹⁹⁶ Because the vagaries of interrogation are even more pronounced, the same holding is imperative in that context.

In *Toward Taping*, I also argued that the constitutional right to a recording is nonwaivable.¹⁹⁷ If it could be waived one can predict that, just as defendants routinely forego their *Miranda* rights, they would often be persuaded to give up their right to a recording.¹⁹⁸ The nonwaivability argument rests on the assumption that recording is vital to determining the reliability (as well as the coerciveness) of interrogations. Consider the fact that a defendant may not waive the right to be tried while competent because society, not just the defendant, has a strong interest in ensuring the integrity of the trial process and a meaningful confrontation between the accused and his or her accusers.¹⁹⁹ The same reasoning supports a nonwaivable right to recording.²⁰⁰

B. Extending the Recording Right Beyond the Stationhouse

One of the more revolutionary developments of modern policing has been the introduction of the police body camera.²⁰¹ While it has been touted primarily as a way of recording and deterring police brutality,²⁰² it could also serve as a *1193 means of ensuring that any encounter before entering the stationhouse, precustody or postcustody, is accurately depicted at later proceedings. Especially in light of the Supreme Court's holdings that postwarning statements which repeat pre-warning statements are not admissible if the prewarning statement was coerced or the police acted in bad faith,²⁰³ a verbatim accounting of police-suspect interaction from custody onward is crucial.

Accordingly, at least in those jurisdictions that have already required police to wear body cameras,²⁰⁴ the constitutional arguments just canvassed should require that the cameras be turned on during all police-suspect confrontations. While the Sixth Amendment applies only to “criminal prosecutions,”²⁰⁵ in *Toward Taping* I argued that, because the Amendment's relevance to the recording issue stems from the right of confrontation rather than the right to counsel, it extends backward to any action that the government describes at trial, just as the admissibility of all hearsay, whether uttered pre- or postcharging, is governed by confrontation analysis.²⁰⁶ The Fifth Amendment, which applies to any “criminal case,”²⁰⁷ also extends back at least to the time of custody, as *Miranda* held.²⁰⁸ And the Due Process Clause applies whether or not a person is a suspect, and thus is not limited by the criminal prosecution or custody thresholds.²⁰⁹

***1194 IV. A WORD ON INTERROGATION OF SUSPECTED TERRORISTS**

In the wake of 9/11, several commentators suggested that the rules restricting interrogation be relaxed when the person interrogated is suspected of engaging in or conspiring to commit a terrorist act.²¹⁰ Of course, if any statements obtained during such interrogations are sought solely for investigatory or intelligence purposes rather than as trial evidence, the Fifth Amendment, which is focused on excluding compelled testimony, is irrelevant.²¹¹ Even if the government wants to obtain admissible statements, however, relaxation of the usual Fifth Amendment rules is not called for in the counter-terrorism context. This is especially so if the HIG technique turns out to be as effective as its progenitors predict.²¹² Even if more aggressive interrogation tactics are thought to be crucial, however, this Article has made clear that interrogators have plenty of second-generation tools at their disposal that fall short of coercion.

If negotiation or first-generation techniques are nonetheless deployed in a terrorism interrogation, the Fifth Amendment or Due Process Clause should require exclusion. Some have argued that the public safety exception to *Miranda*, adopted in *New York v. Quarles*,²¹³ would authorize such tactics.²¹⁴ But the *Quarles* exception only applies in cases of imminent

danger.²¹⁵ And even when it applies, its impact is limited; *Quarles* strongly suggested that statements *1195 resulting from coercion beyond a failure to give *Miranda* warnings are still inadmissible against criminal defendants.²¹⁶

Note, however, that the Fifth Amendment only applies in criminal cases. Preventive detention regimes, such as those associated with military commissions, are considered civil in nature.²¹⁷ Thus, the Fifth Amendment would not require exclusion in proceedings designed to preventively detain enemy combatants. Of course, the Due Process Clause might still require exclusion, but a plausible argument can be made that it does not.²¹⁸ At the same time, if exclusion is not required, any unjustifiable coercion, whether aimed at obtaining evidence admissible in such a proceeding or simply at getting information necessary to prevent an attack, should meet with some other sanction.²¹⁹

CONCLUSION

Miranda was an attempt at giving police clear guidelines about interrogation. Other than its warnings requirement, however, it has not done so. While Court decisions since *Miranda* have clarified a number of peripheral issues, they continue to be vague about the types of interrogation tactics police may use to obtain a confession. Further, the Court has been mum about whether interrogation tactics must be memorialized so that courts have a complete and accurate record of their effect.

This Article has suggested a number of rules to fill these gaps, all of them consistent with the Court's jurisprudence to date. The Fifth Amendment's prohibition on compulsion should bar third degree tactics and explicit and implicit negotiation about legal consequences, and any confession resulting *1196 from such tactics should be excluded, even if it is shown to be reliable. However, manipulative techniques that would not be considered coercive if true—including expressions of sympathy or friendship, suggestions of how one might rationalize a confession, and false evidence plays—violate neither the Fifth Amendment nor the Due Process Clause, and a failure to clear up confusion about rights can be excused on lack-of-state-action grounds, even if the result is an unreliable confession. Confessions should nonetheless be excluded on unreliability grounds under the rules of evidence, if testimony describing laboratory and field research and the results of structured individual evaluations can show that a technique or combination of techniques significantly increased the chances of a false confession, and the prosecution is unable to show the confession includes information that only the perpetrator of the crime is likely to know. To ensure accurate information about the interrogation process is available, recording of all phases of the process should be required under the Due Process Clause, the Fifth Amendment, or the Sixth Amendment's Confrontation Clause. These rules should not be relaxed in national security investigations, although coerced statements might be admissible in noncriminal detention proceedings, as long as some other sanction for unjustifiable coercion is available.

Perhaps research comparing third-generation interrogation techniques to second-generation techniques will convince police departments to move toward interrogation processes that do not require aggressive manipulation. Similarly, perhaps more police departments will, on their own, come to realize the many benefits of recording all interrogations. If not, hopefully the Court will adopt more concrete rules about interrogation tactics and recording well before we reach *Miranda*'s 100th anniversary.

Footnotes

^{a1} Milton Underwood Professor of Law, Vanderbilt University Law School. The author would like to thank the participants in the Symposium honoring the fiftieth anniversary of *Miranda v. Arizona* at Boston University School of Law on September 30 and October 1, 2016, especially Richard Leo and Eve Brensike Primus.

¹ 384 U.S. 436 (1966).

2 *Id.* at 478-79.

3 See CHARLES H. WHITEBREAD & CHRISTOPHER SLOBOGIN, CRIMINAL PROCEDURE: AN ANALYSIS OF CASES AND CONCEPTS 371-96 (6th ed. 2015) (describing Supreme Court case law interpreting *Miranda*).

4 Even police manuals at the time counseled against the third degree. See Brief for ACLU as Amicus Curiae at 4, *Escobedo v. Illinois*, 378 U.S. 478 (No. 615) (stating that police manuals were “not exhibits in a museum of third degree horrors” but rather sought to “carefully advise the police interrogator to avoid tactics which are clearly coercive under prevailing law”).

5 Virtually all commentators agree on this point and have done so for years. See, e.g., William J. Stuntz, *Miranda's Mistake*, 99 MICH. L. REV. 975, 981 n.19 (2001) (“The need for video- and audio-taping is the one proposition that wins universal agreement in the *Miranda* literature.”). This view is expressed by those who oppose *Miranda*, JOSEPH D. GRANO, CONFESSIONS, TRUTH, AND THE LAW 116, 121 (1993); Paul G. Cassell, *Miranda's Social Costs: An Empirical Reassessment*, 90 NW. U. L. REV. 387, 487-91 (1996), as well as by those who support it, YALE KAMISAR, POLICE INTERROGATION AND CONFESSIONS 129-37 (1980); Steven A. Drizin & Richard A. Leo, *The Problem of False Confessions in the Post-DNA World*, 82 N.C. L. REV. 891, 997 (2004).

6 See *infra* text accompanying notes 42-52, 117-26.

7 Several cases, including *Miranda*, have noted that the presence of an attorney can assure an accurate account of the interrogation. *Miranda*, 384 U.S. at 470; see also *Arizona v. Roberson*, 486 U.S. 675, 682 n.4 (1988) (noting that the presence of counsel is a protective feature and one basis for the *Miranda* rule). But the Court has never even suggested that recording might be a good practice.

8 Christopher Slobogin, *Lying and Confessing*, 39 TEX. TECH. L. REV. 1275, 1285-89 (2007).

9 Christopher Slobogin, *Toward Taping*, 1 OHIO ST. J. CRIM. L. 309, 314-21 (2003) (arguing that taping is required by various constitutional provisions).

10 See *infra* note 210.

11 Richard Leo describes the third degree to include “blatant physical abuse,” “deniable physical abuse,” “orchestrated physical abuse,” “incommunicado interrogation,” “food, sleep, and other deprivations,” and “explicit threats of harm.” RICHARD A. LEO, POLICE INTERROGATION AND AMERICAN CRIMINAL JUSTICE 47-54 (2008).

12 *Id.* at 70-74 (stating that as a result of the impact of the Wickersham Report and efforts to professionalize police, by the mid-1960s “the third degree had virtually disappeared”); see also FRANK E. ZIMRING & RICHARD S. FRASE, THE CRIMINAL JUSTICE SYSTEM: MATERIALS ON THE ADMINISTRATION AND REFORM OF THE CRIMINAL LAW 132 (1980) (stating that “today the third degree is almost nonexistent”).

13 See LEO, *supra* note 11, at 109.

14 FRED E. INBAU ET AL., CRIMINAL INTERROGATION AND CONFESSIONS (5th ed. 2013). While the first author is Fred Inbau, Reid (who died in 1982) was the principal developer of the tactics described in the book, and training programs using the book refer to the “Reid Technique.” *Id.* at viii; see also John E. Reid & Associates, Inc., <https://www.reid.com> [<https://perma.cc/WQ7Q-HWYR>] (last visited Jan. 29, 2017).

15 INBAU ET AL., *supra* note 14, at 210-50.

16 See Mariam S. Gohara, *A Lie for a Lie: False Confessions and the Case for Reconsidering the Legality of Deceptive Interrogation Techniques*, 33 FORDHAM URB. L.J. 791, 793-94 (2006) (arguing that empirical evidence shows deception can cause false confessions); Patrick M. McMullen, *Questioning the Questions: The Impermissibility of Police Deception in Interrogations of Juveniles*, 99 NW. U. L. REV. 971, 975 (2005) (finding that deception during interrogation often “effectively leaves the suspect with no rational choice but to confess”); Margaret L. Paris, *Lying to Ourselves*, 76 OR. L. REV. 817, 825, 832 (1997) (arguing that police lying during interrogation is “unjustifiable ... because it is unnecessary and harmful as well as impossible to restrain within reasonable limits” and because it undermines government-citizen trust); Deborah Young, *Unnecessary Evil: Police Lying in Interrogations*, 28 CONN. L. REV. 425, 426 (1996) (stating that police lying during interrogation “actually may

decrease the valid evidence police obtain” and that “[l]ying also reduces integrity in the criminal justice system, a harm that reverberates beyond the individual case” (emphasis omitted)).

17 See, e.g., Young, *supra* note 16, at 475 (“The continued acceptance of police lying is based on the long practice of police deception and an unsubstantiated belief that such lying is necessary for successful prosecutions.”).

18 See Brent Snook et al., *The Next Stage in the Evolution of Interrogations: The PEACE Model*, 18 CAN. CRIM. L. REV. 219, 230-35 (2014) (describing the content of the PEACE model).

19 See HIGH-VALUE INTERROGATION GRP. RES. UNIT, SCIENCE-BASED INTERROGATION: A REFERENCE GUIDE 104-05 (2016) [hereinafter HIG MANUAL].

20 Snook et al., *supra* note 18, at 239.

21 *Id.* at 232 (“If a discrepancy is identified, the interviewer may decide to challenge it at the end of the interview. Challenges are not conducted in an aggressive or accusatorial manner.”).

22 HIG MANUAL, *supra* note 19, at 79-81.

23 *Id.* at 81-82 (“Methods that impose [an increased expenditure of mental energy] serve to increase the amount of cognitive load a person is experiencing. When there is an increase in the level of mental effort liars have to expend, their stories may lack detail and become less logical or coherent.”).

24 *Id.* at 51 (“A common way to start an interview is by building rapport and encouraging cooperation from the subject.”); *id.* at 100-01 (discussing the importance of establishing trust and noting that “[t]here is a fine line between trust and honesty”); Snook et al., *supra* note 18, at 231 (describing the importance of rapport-building and creating a “working alliance”).

25 HIG MANUAL, *supra* note 19, at 82-89.

26 Simon Oleszkiewicz et al., *The Scharff-Technique: Eliciting Intelligence from Human Sources*, 38 LAW & HUM. BEHAV. 478, 479 (2014) (pointing out that this illusion leads the suspect to believe that the interrogator is “very well informed on the topic”).

27 Cf. Daniel W. Shuman, *The Use of Empathy in Forensic Examinations*, 3 ETHICS & BEHAV. 289, 294-97 (1993) (arguing that evincing empathy during a forensic interview is often unethical).

28 Snook et al., *supra* note 18, at 236 (stating that the confession rate is about fifty percent whether the Reid or PEACE technique is used); see also GISLI H. GUDJONNISON, *THE PSYCHOLOGY OF INTERROGATIONS AND CONFESSIONS: A HANDBOOK* 620-23 (2003) (stating, based on experience in the United Kingdom, that fears that abandoning the Reid technique would reduce confessions “may be overstated”).

29 First, confession rates are probably higher in the United States, where the Reid technique is prevalent, than in the United Kingdom. See Slobogin, *supra* note 8, at 1282-83 nn.44-45. Second, United Kingdom police still use Reid-type techniques in some cases. *Id.* Third, United Kingdom confession rates are probably enhanced by the fact that United Kingdom interrogators routinely tell suspects who refuse to answer a question that “[i]t may harm your defence if you do not mention when questioned something which you later rely on in Court,” and by knowledge on the part of the interrogators that fruits of a confession will not be excluded from evidence. Jason Mazzone, *Silence, Self-Incrimination, and Hazards of Globalization*, in *COMPARATIVE CRIMINAL PROCEDURE* 308, 321 (Jacqueline E. Ross & Stephen C. Thaman eds., 2016).

30 See HIG MANUAL, *supra* note 19, at 23 (describing research finding that “[m]otivational [i]nterviewing,” a variant of nonaccusatorial questioning, “increased Information Yield both directly and indirectly”); Oleszkiewicz et al., *supra* note 26, at 482 (finding that the Scharff technique “resulted in significantly more new information than the Direct Approach”).

31 HIG MANUAL, *supra* note 19, at 80.

32 This conclusion follows from the false confession research, which indicates that, outside of interrogations involving very young children or people with intellectual disability, most false confessions come in cases using Reid-type techniques or the third degree. See *infra* notes 34 & 158 and text accompanying note 139.

- 33 According to one source, “systematic study” assessing the effectiveness of the PEACE method “in the field” has yet to be completed. Snook et al., *supra* note 18, at 233. *But see* S. Soukara et al., *What Really Happens in Police Interviews with Suspects? Tactics and Confessions*, 15 PSYCHOL. CRIME & L. 493, 504 (2009).
- 34 *See, e.g.*, DAVID DIXON (WITH GAIL TRAVIS), INTERROGATING IMAGES: AUDIO-VISUALLY RECORDING POLICE QUESTIONING OF SUSPECTS 228-29 (2009) (reporting survey of police, prosecutors, defense attorneys, and judges in England indicating that the first three groups were more likely to state that the PEACE technique had decreased the number of confessions, although also indicating that all four groups believed it increased guilty pleas); Jennifer T. Perillo & Saul M. Kassin, *Inside Interrogation: The Lie, the Bluff, and False Confessions*, 35 LAW & HUM. BEHAV. 327, 334 (2011) (reporting a laboratory study finding that while the bluff technique increased false confessions from forty-five percent to seventy percent it also increased true confessions from twenty-six percent to eighty-nine percent); Melissa Russano et al., *Investigating True and False Confessions Within a Novel Experimental Paradigm*, 16 PSYCHOL. SCI. 481, 484 tbl.1 (2005) (reporting a laboratory study finding that manipulative techniques, while increasing false confessions from 6% to 18%, increased true confessions from 46% to 81%). The one laboratory study that purports to compare second- and third-generation techniques found that the two techniques produced roughly the same percentage of true confessions. Fadia M. Narchet et al., *Modeling the Influence of Investigator Bias on the Elicitation of True and False Confessions*, 35 LAW & HUM. BEHAV. 452, 459 tbl.3 (2011). But it appears that the interrogators in this study switched to Reid techniques if they could not get a confession using “non-coercive” techniques, *see id.* at 458, which would reduce the denominator of the confession rate for the latter techniques. Relatedly, field research indicates that the use of Reid techniques increases the probability of a confession (although here it is not known whether the confession is true or false). *See, e.g.*, Lesley King & Brent Snook, *Peering Inside a Canadian Interrogation Room: An Examination of the Reid Model of Interrogation, Influence Tactics, and Coercive Strategies*, 36 CRIM. JUST. & BEHAV. 674, 690 (2007); Soukara et al., *supra* note 33, at 503.
- 35 *See infra* text accompanying notes 173-76.
- 36 479 U.S. 157, 167 (1986).
- 37 *Rogers v. Richmond*, 365 U.S. 534, 540 (1961) (“Our decisions under [the Due Process Clause] have made clear that convictions following the admission into evidence of confessions which are involuntary, i.e., the *product* of coercion, either physical or psychological, cannot stand.”). For further discussion of the meaning of “involuntary” under the Due Process Clause, *see infra* text accompanying notes 112-26.
- 38 Stephen J. Schulhofer, *Miranda, Dickerson, and the Puzzling Persistence of Fifth Amendment Exceptionalism*, 99 MICH. L. REV. 941, 950 (2001) (“*Miranda* ... brought Fifth Amendment standards into the stationhouse under the expressly stated assumption that those standards provided more protection than the traditional Fourteenth Amendment voluntariness requirement.”).
- 39 *Miranda v. Arizona*, 384 U.S. 436, 476 (“The requirement of warnings and waiver of rights is a fundamental with respect to the Fifth Amendment privilege and not simply a preliminary ritual to existing methods of interrogation.”).
- 40 *Id.* at 444 (“The defendant may waive effectuation of these rights, provided the waiver is made voluntarily, knowingly and intelligently.”). The Court has since downplayed the last two requirements. *See infra* text accompanying notes 89-97.
- 41 *See, e.g.*, *Miranda*, 384 U.S. at 505 (Harlan, J., dissenting) (“I believe that reasoned examination will show that the Due Process Clauses provide an adequate tool for coping with confessions and that, even if the Fifth Amendment privilege against self-incrimination be invoked, its precedents taken as a whole do not sustain the present rules.”); Albert W. Alschuler, *Miranda's Fourfold Failure*, 97 B.U. L. REV. 851, 892-93 (2017).
- 42 *See generally* Michael Kates, *Markets, Sweatshops and Coercion*, 13 GEO. J.L. & PUB. POL'Y 367, 368 (2015) (“Coercion is a philosophically contested concept. Indeed, the problem is even worse than that. For not only is there sharp disagreement in the philosophical literature as to what is the correct definition or meaning of coercion but the nature of that disagreement ranges over a number of different dimensions as well.”).
- 43 *See, e.g.*, *Beecher v. Alabama*, 389 U.S. 35, 36 (1967) (finding coercion where police officers held a gun to the head of a wounded confessant in a successful effort to extract a confession); *Davis v. North Carolina*, 384 U.S. 737, 746-51 (1966) (finding that sixteen days of incommunicado interrogation in a closed cell without windows, limited food, and coercive tactics

constituted coercion); *Reck v. Pate*, 367 U.S. 433, 435-39 (1961) (finding coercion where a defendant was held for four days with inadequate food and medical attention until a confession was obtained); *Culombe v. Connecticut*, 367 U.S. 568, 570 (1961) (finding coercion where a defendant held for five days of repeated questioning during which police employed coercive tactics); *Payne v. Arkansas*, 356 U.S. 560, 567 (1958) (finding coercion where defendant held incommunicado for three days with little food and told that Chief of Police was preparing to admit lynch mob into jail); *Ashcraft v. Tennessee*, 322 U.S. 143, 149 (1944) (finding coercion where defendant was questioned by relays of officers for thirty-six hours without an opportunity for sleep).

44 *Lynumn v. Illinois*, 372 U.S. 528, 532 (1963) (holding invalid a confession that female suspect said she made “because the police told me they were going to send me to jail for 10 years and take my children, and I would never see them again; so I agreed to say whatever they wanted me to say”); *Rogers v. Richmond*, 365 U.S. 534, 536 (1961) (holding invalid a confession obtained after police threatened to arrest suspect's wife).

45 See *Miranda*, 384 U.S. at 476 (“[A]ny evidence that the accused was threatened, tricked, or cajoled into a waiver will, of course, show that the defendant did not voluntarily waive his privilege.”).

46 See *Spano v. New York*, 360 U.S. 315, 323 (1959) (involving a police officer, a friend of Spano's, falsely stating he would lose his job and that his family would suffer if a confession was not forthcoming); *Leyra v. Denno*, 347 U.S. 556, 559-60 (1954) (involving a psychiatrist posing as a medical doctor who repeatedly told Leyra “how much he wanted to and could help him, how bad it would be for petitioner if he did not confess, and how much better he would feel, and how much lighter and easier it would be on him if he would just unbosom himself to the doctor”).

47 See *Oregon v. Mathiason*, 429 U.S. 492, 495-96 (1977) (suggesting that lying about finding the suspect's fingerprints at the scene of the crime was not relevant to the admissibility issue); *Frazier v. Cupp*, 394 U.S. 731, 739 (1969) (holding admissible a confession by a suspect who was told, falsely, that his codefendant had just confessed).

48 See, e.g., *Miller v. Fenton*, 796 F.2d 598, 602 (3d Cir. 1986) (upholding confession obtained after the detective stated several times “I'm your brother.”).

49 See *id.* at 603 (“I know how you feel inside, Frank, it's eating you up, am I right? It's eating you up, Frank. You've got to come forward. You've got to do it for yourself, for your family, for your father, this is what's important, the truth, Frank.”); *United States v. Huggins-McLean*, No. CR414-141, 2015 WL 370237, at *4 (S.D. Ga. Jan. 28, 2015) (upholding confession obtained after police told the suspect “it would be better for him to speak and provide a ‘truthful and honest’ statement about his criminal activities”).

50 Paul Marcus, *It's Not Just About Miranda: Determining the Voluntariness of Confessions in Criminal Prosecutions*, 40 VAL. U. L. REV. 601, 612-13 (2006) (noting that courts have permitted lies about: “witnesses against the defendant, earlier statements by a now-deceased victim, an accomplice's willingness to testify, whether the victim had survived an assault, ‘scientific’ evidence available, including DNA and fingerprint evidence, and the degree to which the investigating officer identified and sympathized with the defendant”).

51 *Id.* at 623 (“[T]here are many cases in which confessions are found to be voluntary based upon a variety of promises made, including vague guarantees that the defendant will receive better treatment if she confesses, offers of more lenient punishment for the suspect, assurances of lesser charges being prosecuted if the individual confesses” (footnotes omitted)).

52 Margaret L. Paris, *Trust, Lies, and Interrogation*, 3 VA. J. SOC. POL'Y & L. 3, 6 (1995) (“The Court's few pronouncements in this area have been so enigmatic, and so highly contingent on specific facts, that they are largely ignored by interrogators and courts alike.”).

53 Slobogin, *supra* note 8, at 1285-89.

54 *Id.* at 1287. I also argued, based on the work of moral philosopher Sissela Bok and Fourth Amendment jurisprudence, that deception should only be permissible when directed at people for whom police have probable cause. *Id.* at 1276-80; see also SISSELA BOK, LYING: MORAL CHOICE IN PUBLIC AND PRIVATE LIFE (1978)). Thus, interrogation prior to arresting an individual would be impermissible. See Christopher Slobogin, *Deceit, Pretext, and Trickery: Investigative Lies by the Police*, 76 OR. L. REV. 775, 811 (1997).

- 55 Slobogin, *supra* note 8, at 1285-86 (outlining three alternative schools of thought about coercion and asserting that all three regard such practices as violations).
- 56 Examples of cases holding that the Fifth Amendment prohibits such penalties include *Lefkowitz v. Cunningham*, 431 U.S. 801, 803-04 (1977) (loss of the right to participate in political associations and to hold public office), *Lefkowitz v. Turley*, 414 U.S. 70, 73-74 (1973) (ineligibility to receive government contracts), *Uniformed Sanitation Men Ass'n v. Comm'r of Sanitation*, 392 U.S. 280, 281 (1968) (termination of employment), and *Spevack v. Klein*, 385 U.S. 511, 512-13, 516 (1967) (loss of professional license). *But see McKune v. Lile*, 536 U.S. 24, 29, 48 (2002) (concluding that the Fifth Amendment is not violated by conditioning the extent to which prisoners can keep personal property, see visitors, spend money, and earn money on self-incrimination).
- 57 *See* DAVID SIMON, *HOMICIDE: A YEAR ON THE KILLING STREETS* 194-95 (1991) (stating that Baltimore detectives routinely tell a suspect that an invocation of rights will “make matters worse for him, for it would prevent his friend the detective, from writing up the case as manslaughter or perhaps even self-defense, rather than first degree murder”).
- 58 *See, e.g., Rose v. Lee*, 252 F.3d 676, 686 (4th Cir. 2001) (“[W]e decline to hold that the cryptic promise that ‘things would go easier’ on [the suspect] if he confessed amounts to unconstitutional coercion.”); LEO, *supra* note 11, at 158 (describing cases where interrogators promised lighter sentences if the suspect confessed).
- 59 It is well known that sentences imposed after guilty pleas are often much shorter than those that would have been imposed after trial. *See, e.g., Nancy J. King et al., When Process Affects Punishment: Differences in Sentences After Guilty Plea, Bench Trial, and Jury Trial in Five Guidelines States*, 105 COLUM L. REV. 959, 992, 1005-09 (2005) (finding increases in sentences for those who go to trial ranging “from 13% to 461% in Washington, from 58% to 349% in Maryland, and from 23% to 95% in Pennsylvania”).
- 60 *See* INBAU ET AL., *supra* note 14, at 344 (stating that police should not tell a suspect who is denying the crime “I will not only charge you with this offense but also with obstruction of justice”); *id.* at 345 (stating that interrogators should not tell a suspect “if this is the first time you did something like this, I’ll talk to the judge and make sure that he gives you probation”).
- 61 *See id.* at 345 (stating that an interrogator may say to a suspect “if this is something that happened on the spur of the moment, that would be important to include in my report”); *see also id.* at 296, 299 (recommending that, at the climactic stage of the interrogation, the suspect who continues to deny the crime be given only two alternatives--e.g., “If you’ve done this dozens of times before, that’s one thing. But if this was just the first time it happened, that would be important to establish” or “Joe, this is very critical. When you pulled that trigger were you just trying to slightly injure him or were you aiming for his heart?”).
- 62 *See* LEO, *supra* note 11, at 132 (providing examples).
- 63 *See* INBAU ET AL., *supra* note 14, at 296 (discussing and approving these techniques).
- 64 Tough cases illustrate the thin line between negotiation and rationalization. *See, e.g., Miller v. Fenton*, 796 F.2d 598, 609 (3d Cir. 1986) (finding no coercion where officer, posing as someone who wanted to help the suspect, stated “[y]ou are not responsible” and “[y]ou are not a criminal,” because detective “never stated that anyone but he thought that Miller was ‘not a criminal,’ nor did he state that he had any authority to affect the charges brought against Miller”); *Fundaro v. Curtin*, No. 4:13-cv-11868, 2015 WL 357012, at *7 (E.D. Mich. Jan. 26, 2015) (finding no coercion when police statements “were conditional: *if* Petitioner acted in self-defense, *then* he should explain his side of the story. The statements did not inform him that he in fact acted in self-defense.”); *People v. Holloway*, 91 P.3d 164, 178 (Cal. 2004) (holding that the detective’s “general assertion that the circumstances of a killing could ‘make[] a lot of difference’ to the punishment” did not invalidate the confession) (alteration in original). In *Fundaro* and *Holloway*, the police were only suggesting the circumstances under which lenient treatment would occur, not that lenient treatment would be forthcoming if the suspect confessed. In *Miller*, the officer, in a part of the transcript that the court does not discuss, tied the suspect’s entitlement to help to a confession. *Miller*, 796 F.2d at 623, 638 (stating repeatedly, “I can’t help you without the truth”). Thus, the latter officer came much closer to suggesting leniency in exchange for a confession, and coercion should have been found.
- 65 168 U.S. 532 (1897).
- 66 *Id.* at 542-43.

- 67 397 U.S. 742 (1970).
- 68 *Id.* at 754.
- 69 429 U.S. 28 (1976).
- 70 *Id.* at 30 (admitting confession solicited *after* a bargain but reaffirming *Bram*).
- 71 499 U.S. 279 (1991).
- 72 *Id.* at 285 (citing *Bram*, 168 U.S. at 542-43).
- 73 *See* Marcus, *supra* note 50, at 621-22).
- 74 *See* Alschuler, *supra* note 41, at 865 (“When our justice system does not balk at using promises of leniency to induce the ultimate act of self-incrimination--a plea of guilty--it need not be squeamish about using similar leverage to induce suspects to say truthfully what happened.”); Lawrence Rosenthal, *Against Orthodoxy: Miranda Is Not Prophylactic and the Constitution Is Not Perfect*, 10 CHAP. L. REV. 579, 600-01 (2007) (“[U]nder the guilty-plea cases, even if the accused and his counsel misapprehend the strength of the prosecution’s case or the availability of defenses, a guilty plea is still considered a valid waiver. A *Miranda* waiver is certainly no less valid if the suspect somehow misapprehends his own best interests.”).
- 75 *Santobello v. New York*, 404 U.S. 257, 260 (1971) (stating that plea negotiation “is an essential component of the administration of justice” and that “[p]roperly administered, it is to be encouraged”).
- 76 *Bordenkircher v. Hayes*, 434 U.S. 357, 363 (1978) (citing *Brady v. United States*, 397 U.S. 742, 758 (1970)).
- 77 *See* LEO, *supra* note 11, at 29 (using the term “pre-plea negotiation”); *id.* at 133 (stating that interrogators seek to convince the suspect that “[his] admission is, in effect, his quid pro quo for an end to the interrogation and avoidance of the worse-case scenario--harsher treatment or punishment, for example”).
- 78 *See supra* note 47; *infra* note 80.
- 79 Welsh S. White, *Police Trickery in Inducing Confessions*, 127 U. PA. L. REV. 581, 614-15 (1979).
- 80 *See, e.g., Illinois v. Perkins*, 496 U.S. 292, 294 (1990) (holding that questioning by an undercover officer posing as a jail inmate does not violate Fifth Amendment). What if the interrogator poses as a lawyer? Although there would be no coercion, the warnings about silence and counsel would have to be given, which makes this scenario practically impossible.
- 81 William J. Stuntz, *Waiving Rights in Criminal Procedure*, 75 VA. L. REV. 761, 823 (1989) (“[Deception] avoids the confession-or-perjury dilemma either by convincing the suspect that truthful statements will not have incriminating consequences, or by making him forget temporarily that they will.”).
- 82 *See Miranda v. Arizona*, 384 U.S. 436, 452-53 (1966) (describing the technique).
- 83 *See* Richard J. Ofshe & Richard A. Leo, *The Decision to Confess Falsely: Rational Choice and Irrational Action*, 74 DEN. U. L. REV. 979, 1056-60, 1077 (1997) (providing these examples and distinguishing between the moral and psychological consequences of not confessing and the legal consequences of not doing so); INBAU ET AL., *supra* note 14, at 289 (suggesting, *inter alia*, that suspects be told to tell the truth “for the sake of everyone concerned”).
- 84 *See supra* note 64. In one of the first studies to investigate interrogation techniques, participants perceived explicit threats and promises to be more coercive than indirect maximization and minimization techniques, which tended to be seen as no more coercive than simple questioning. Saul M. Kassin & Karlyn McNall, *Police Interrogations and Confessions: Communicating Promises and Threats by Pragmatic Implication*, 15 LAW & HUM. BEHAV. 233, 238 (1991). In a separate experiment, participants were much more likely to find that minimization techniques that were combined with the statement “If you just tell me the truth, we can get this matter straightened out” were a form of negotiation than when the latter sentence was not included. *See id.* at 240-41 (describing three scenarios--one without the statement and two with the statement).

- 85 See ARTHUR S. AUBRY, JR. & RUDOLPH R. CAPUTO, CRIMINAL INTERROGATION 85-86 (1965) (recommending “bluffing” the suspect by telling him, e.g., that he was seen at the scene of the crime or that the co-defendant has confessed).
- 86 See, e.g., *State v. Cayward*, 552 So. 2d 971, 972 (Fla. Dist. Ct. App. 1989) (excluding confession produced after police showed suspect fabricated report showing that the semen stains on the victim's underwear belonged to the suspect); *Washoe Cty. v. Bessey*, 914 P.2d 618, 619, 622 (Nev. 1996) (refusing to exclude confession on similar facts); *State v. Patton*, 826 A.2d 783, 784 (N.J. Super. Ct. App. Div. 2003) (prohibiting use of a manufactured recording in which police presented a fictitious eyewitness).
- 87 *Cayward*, 552 So. 2d at 974 (“A report falsified for interrogation purposes might well be retained and filed in police paperwork. Such reports have the potential of finding their way into the courtroom.”).
- 88 See George C. Thomas III, *Regulating Police Deception During Interrogation*, 39 TEXAS TECH. L. REV. 1293, 1308-12, 1316-19 (2007) (describing *Cayward*, *Bessey*, and *Payton* and concluding that the confessions should have been admitted).
- 89 See, e.g., *Hart v. Attorney Gen. of Fla.*, 323 F.3d 884, 894-95 (11th Cir. 2003) (excluding a statement from the suspect who was told “honesty will not hurt you”); *Ex parte Johnson*, 522 So. 2d 234, 235-36 (Ala. 1988) (excluding a statement after the suspect was told it could not be used in a criminal case); *Commonwealth v. Peters*, 373 A.2d 1055, 1058-63 (Pa. 1977) (excluding statements made after the suspect was told they would only be used against other suspects); *State v. Stanga*, 617 N.W.2d 486, 490-91 (S.D. 2000) (excluding a statement from a suspect who was told the statements would just be “between the two of them”).
- 90 See *supra* text accompanying note 40.
- 91 441 U.S. 369, 371 (1979) (involving a suspect who, after receiving the warnings, stated “I will talk to you but I am not signing any form”).
- 92 479 U.S. 523, 525 (1987) (involving a suspect who said he understood his rights and then said he would not give a written statement without a lawyer being present but had “no problem” talking).
- 93 560 U.S. 370 (2010).
- 94 *Id.* at 375-76 (involving a suspect who was given warnings and said very little for the first two hours and forty-five minutes of the interrogation, at which point, in answer to the question “Do you pray to God to forgive you for shooting that boy down?” he answered “Yes”).
- 95 512 U.S. 452, 455 (1994) (holding that police may continue questioning a suspect who states “Maybe I should talk to a lawyer” without stopping to clarify if the suspect wants counsel).
- 96 *Id.* at 455 (holding that police may continue questioning a suspect who states “Maybe I should talk to a lawyer” without stopping to clarify if the suspect wants counsel).
- 97 479 U.S. 564 (1987).
- 98 *Id.* at 567 (where suspect waived his rights on the understanding police would question him about a firearms charge, and at some later point during the ninety-minute interview answered affirmatively when asked if he had ever shot someone).
- 99 See, e.g., *Berghuis*, 560 U.S. at 404 (Sotomayor, J., dissenting) (“Today's decision bodes poorly for the fundamental principles that *Miranda* protects.”); *Davis*, 512 U.S. at 472 (Souter, J., concurring) (“When a suspect understands his (expressed) wishes to have been ignored (and by hypothesis, he has said something that an objective listener could ‘reasonably,’ although not necessarily, take to be a request), in contravention of the ‘rights’ just read to him by his interrogator, he may well see further objection as futile and confession (true or not) as the only way to end his interrogation.”); *Spring*, 479 U.S. at 579 (Marshall, J., dissenting) (“[R]equiring the officers to articulate at a minimum the crime or crimes for which the suspect has been arrested could contribute significantly toward ensuring that the arrest was in fact lawful and the suspect's statement not compelled because of an error at this stage alone.”); Yale Kamisar, *The Rise, Decline and Fall (?) of Miranda*, 87 WASH. L. REV. 965, 1008-20 (2012) (criticizing *Butler*, *Berghuis*, and *Davis*).

- 100 [Colorado v. Connelly](#), 479 U.S. 157, 170 (1986).
- 101 See Robert P. Mosteller, *Police Deception Before Miranda Warnings: The Case for Per Se Prohibition of an Entirely Unjustified Practice at the Most Critical Moment*, 39 TEX. TECH L. REV. 1239, 1270 (2007) (arguing that “passive deception” at the time of the warnings is permissible, but that “affirmative, false statements” should not be).
- 102 See *supra* note 97.
- 103 See Richard Rogers et al., *An Analysis of Miranda Warnings and Waivers: Comprehension and Coverage*, 31 LAW & HUM. BEHAV. 177, 188-91 (2007) (finding that understanding the warnings requires a reading capability somewhere between sixth and tenth grade, which many defendants lack); Richard Rogers et al., “*Everyone Knows Their Miranda Rights*”: *Implicit Assumptions and Countervailing Evidence*, 16 PSYCHOL. PUB. POLY & L. 300, 307-11 (2010) (finding, among a sample of adult defendant and college students given the warnings, that 30.2% believed that once counsel is requested questioning may continue until counsel arrives, 30% believed that silence could be used as evidence, 25.9% believed that a waiver must be signed to be valid, and 12.8% believed that statements could be retracted).
- 104 In *Davis*, for instance, the defendant later unequivocally stated he did not want counsel. *Davis*, 512 U.S. at 455. In *Spring*, the suspect not only was given the usual warnings but was also told he had the right to cut off questioning at any time. *Spring*, 479 U.S. at 564. In *Berghuis*, the suspect never said he wanted to remain silent and in fact responded to a number of the interrogators' questions. *Thompkins*, 560 U.S. at 375.
- 105 See LEO, *supra* note 11, at 125, 127 (describing such ruses).
- 106 Morgan Cloud et al., *Words Without Meaning: The Constitution, Confessions, and Mentally Retarded Suspects*, 69 U. CHI. L. REV. 495, 513, 576 (2002) (finding extremely low understanding of the *Miranda* warnings among people with an IQ below eighty and noting that “[n]umerous participants involved in our study [of interrogation] answered yes to questions that they neither understood nor were able to answer”); Carol K. Sigelman et al., *When in Doubt, Say Yes: Acquiescence in Interviews with Mentally Retarded Persons*, 19 MENTAL RETARDATION 53, 53-57 (1981) (discussing studies that show individuals with intellectual disability are more likely to comply with unreasonable instructions); Susan Harter, *Mental Age, IQ, and Motivational Factors in the Discrimination Learning Set Performance of Normal and Retarded Children*, 5 J. EXPERIMENTAL CHILD PSYCHOL. 123, 137-38 (1967) (finding that individuals with intellectual disability seek approval from authority figures even when it requires giving an answer they know to be incorrect); Kimberly Larson, *Improving the “Kangaroo Courts”: A Proposal for Reform in Evaluating Juveniles' Waiver of Miranda*, 48 VILL. L. REV. 629, 657 (2003) (summarizing psychological research reporting that “children are more compliant and suggestible than adults”); Amye R. Warren & Dorothy F. Marsil, *Why Children's Suggestibility Remains a Serious Concern*, 65 LAW & CONTEMP. PROBS. 127, 128-31 (2002) (summarizing research indicating that children under twelve are significantly more suggestible than older children and adults).
- 107 See generally THOMAS GRISSO, EVALUATING COMPETENCIES: FORENSIC ASSESSMENTS AND INSTRUMENTS 149-92 (2003) (describing methods for evaluating waivers of rights by, inter alia, juveniles and people with intellectual disability).
- 108 See, e.g., *State v. Turner*, 847 N.W.2d 69, 73-74 (Neb. 2014) (refusing to exclude confession despite use of negotiation tactic--misrepresenting that a lesser sentence would be imposed for felony murder-- because the confession was immediately preceded by rationalization techniques-- telling suspect he was not an evil person, exhorting him to “do the right thing,” and discussing the fate of his soul).
- 109 [Lego v. Twomey](#), 404 U.S. 477, 489 (1972).
- 110 [Michigan v. Harvey](#), 494 U.S. 344, 351 (1990) (“We have mandated the exclusion of reliable and probative evidence for all purposes ... when it is derived from involuntary statements.” (citing [New Jersey v. Portash](#), 440 U.S. 450, 459 (1979))).
- 111 [Colorado v. Connelly](#), 479 U.S. 157, 170 (1986) (“The sole concern of the Fifth Amendment, on which *Miranda* was based, is governmental coercion.”).
- 112 The strongest such statement came in [Lisenba v. California](#), 314 U.S. 219, 236 (1941) (“The aim of the rule that a confession is inadmissible unless it was voluntarily made is to exclude false evidence.”). But most of the Court's due process cases during the

pre-*Miranda* era simply emphasized that the focus should be coercion, not reliability. See [Rogers v. Richmond](#), 365 U.S. 534, 540-41 (1961) (“Our decisions ... have made clear that convictions following the admission into evidence of confessions which are involuntary, *i.e.*, the product of coercion, either physical or psychological, cannot stand ... not because such confessions are unlikely to be true but because the methods used to extract them offend an underlying principle in the enforcement of our criminal law: that ours is an accusatorial and not an inquisitorial system--a system in which the State must establish guilt by evidence independently and freely secured and may not by coercion prove its charge against an accused out of his own mouth.”).

113 *Connelly*, 497 U.S. at 163-67 (holding that introducing respondent's statements into evidence did not constitute a violation of the Due Process Clause).

114 *Id.* at 162.

115 *Id.* at 167.

116 *Id.*

117 The Court summarized the relevant sentiment in [Miller v. Fenton](#), 474 U.S. 104, 109 (1985) (“This Court has long held that certain interrogation techniques, either in isolation or as applied to the unique characteristics of a particular suspect, are so offensive to a civilized system of justice that they must be condemned under the Due Process Clause of the Fourteenth Amendment.”).

118 *Brown v. Mississippi*, 267 U.S. 278, 286 (1936).

119 *Chambers v. Florida*, 309 U.S. 227, 236 (1940).

120 *Chavez v. Martinez*, 538 U.S. 760, 774 (2003) (quoting [Rochin v. California](#), 342 U.S. 165, 172, 174 (1952)).

121 This point has been made even by commentators who generally want to restrict interrogation practices. See Yale Kamisar, *On the Fortieth Anniversary of the Miranda Case: Why We Needed It, How We Got It--And What Happened to It*, 5 OHIO ST. J. CRIM. L. 163, 168 (2007) (calling the due process involuntariness test “too amorphous, too perplexing, too subjective and too time-consuming to administer effectively”).

122 Jane R. Bambauer & Toni M. Massaro, *Outrageous and Irrational*, 100 MINN. L. REV. 281, 335 (2015) (“In short, conscience-shocking behavior happens, but courts only rarely call it unconstitutional.”).

123 See, e.g., [Arizona v. Fulminante](#), 499 U.S. 279, 285-88 (1991) (where the Court's due process analysis focused solely on “coercion,” and did not mention offensiveness); [Mincey v. Arizona](#), 437 U.S. 385, 398-402 (1978) (where the Court's due process analysis focused on whether police conduct, which included denying multiple requests for an attorney from a hospitalized individual, undermined the suspect's “free and rational choice”). In *Chavez v. Martinez*, the Court explained its holding that questioning of a hospitalized individual did not “shock the conscience” on the ground that the questioning was not “intended to injure in some way unjustifiable by any government interest.” *Chavez*, 538 U.S. at 774-75. Assuming first-generation tactics are not at issue, interrogation is rarely intended to injure a suspect unjustified by a legitimate government interest.

124 E.g., Bambauer & Massaro, *supra* note 122, at 347-48 (arguing that [Missouri v. Seibert](#), 542 U.S. 600 (2004)--where police knowingly took advantage of Supreme Court case law finding admissible postwarning statements made after prewarning statements--should have been decided on due process “outrageousness” grounds rather than on the assumption that the second confession was heavily influenced by the first (an assumption that Court had been unwilling to make in previous cases)). The fact that the Court chose the path it did suggests its antipathy toward using substantive due process in this context.

125 This phrase first appeared in [Oregon v. Elstad](#), 470 U.S. 298, 312 (1985), as a way of distinguishing coercion that renders a confession involuntary under the Due Process Clause and coercion associated with failing to give the warnings. See [Colorado v. Spring](#), 479 U.S. 564, 574 (1987).

- 126 Of course, if a person is not in custody at the time of questioning, see *Fulminante*, 499 U.S. at 283-84, or if the person is in custody but an exception to *Miranda* applies, see *New York v. Quarles*, 467 U.S. 649, 650 (1984), then the Due Process Clause alone protects individuals from coercion during interrogation.
- 127 Eve Primus has argued that, in context, *Connelly* was merely stating that reliability, standing alone, is not guaranteed by the Due Process Clause, and that unreliability caused by state action is still a matter of concern under the Clause. Eve Brensike Primus, *The Future of Confession Law: Toward Rules for the Voluntariness Test*, 114 MICH. L. REV. 1, 32-34 (2015). But the Court's statement in *Connelly* that the Clause is only triggered by "coercion brought to bear on the defendant by the State" forecloses that argument, as does its declaration that "the voluntariness determination has nothing to do with the reliability of jury verdicts; rather, it is designed to determine the presence of police coercion." *Colorado v. Connelly*, 479 U.S. 157, 167-68 (1986); see also BRANDON L. GARRETT, CONVICTING THE INNOCENT: WHERE CRIMINAL PROSECUTIONS GO WRONG 37 (2011) ("The U.S. Supreme Court has held that unreliability is irrelevant to the question whether a confession statement is sufficiently voluntary to be admitted at trial.").
- 128 *Connelly*, 479 U.S. at 167. *Connelly* specifically referenced [Federal Rule of Evidence 601](#), which states that the competency of witnesses is presumed. *Id.* That rule, at best, is tangential to the central inquiry, for reasons developed in the rest of this Section.
- 129 See, e.g., [FED. R. EVID. 401](#) (defining as relevant any evidence that tends to make the existence of any material fact "more or less probable than it would be without the evidence"); [FED. R. EVID. 403](#) (excluding relevant evidence when its probative value is "substantially outweighed by a danger of ... unfair prejudice ... [or] misleading the jury").
- 130 See *Fulminante*, 499 U.S. at 292 ("A defendant's confession is 'probably the most probative and damaging evidence that can be admitted against him'" (quoting *Cruz v. New York*, 481 U.S. 186, 195 (1987))). Certainly, confessions have a profound impact on the jury, so much so that we may justifiably doubt its ability to put them out of mind "even if told to do so." *Id.*; Saul M. Kassin & Katherine Neumann, *On the Power of Confession Evidence: An Experimental Test of the Fundamental Difference Hypothesis*, 21 LAW & HUM. BEHAV. 469, 479, 481 (1997) (finding that confessions are more prejudicial to the defendant's case than eyewitness identification and character testimony).
- 131 See Richard A. Leo et al., *Bringing Reliability Back In: False Confessions and Legal Safeguards in the Twenty-First Century*, 2006 WIS. L. REV. 479, 531-33 ("Because juries often see confession evidence as dispositive of guilt, even when it is false, its prejudicial effect can be devastating to an innocent defendant.").
- 132 Sharon L. Davies, *The Reality of False Confessions--Lessons of the Central Park Jogger Case*, 30 N.Y.U. REV. L. & SOC. CHANGE 209, 232-33 (2006).
- 133 Eugene R. Milhizer, *Confessions After Connelly: An Evidentiary Solution for Excluding Unreliable Confessions*, 81 TEMP. L. REV. 1, 41-47 (2008) (criticizing the *corpus delicti* rule--requiring independent evidence that a crime has occurred--as tangential to the goal of assuring reliable confessions, and the "trustworthiness" rule--requiring corroboration from virtually any source--as too "permissive" toward the prosecution).
- 134 *Id.* at 47-54 (proposing and elaborating on a new rule of evidence governing admissibility of confessions).
- 135 See Leo et al., *supra* note 131, at 531.
- 136 I prefer Rule 401/403 analysis because it uses existing law to confront directly the balance between the State's interest in introducing relevant evidence and the defendant's interest in keeping tainted, highly influential evidence from getting to the factfinder.
- 137 Primus, *supra* note 127, at 41.
- 138 Nor is such proof needed with respect to decisions about the coercion issue, even though state action is required in that setting. Nowhere does *Connelly* require knowledge of wrongdoing; it merely requires coercion. State action doesn't have to be intentional, it just has to exist.
- 139 The research findings bolstering these claims about false confessions are summarized in Saul M. Kassin et al., *Police-Induced Confessions: Risk Factors and Recommendations*, 34 LAW & HUM. BEHAV. 3, 16-22 (2010).

- 140 *Id.*
- 141 *See id.* at 15-16.
- 142 *See supra* text accompanying notes 56-77, 106-07.
- 143 509 U.S. 579, 585-97 (1993).
- 144 *See, e.g.*, FED. R. EVID. 104(a) (stating that in deciding the “preliminary question” about whether evidence is admissible, the “court is not bound by evidence rules”).
- 145 *See, e.g.*, *United States, v. Belyea*, 159 F. App'x 525, 529-30 (4th Cir. 2005) (requiring a particularized *Daubert* inquiry with regards to the admission of expert testimony on false confessions); *People v. Kowalski*, 821 N.W.2d 14, 24-26 (Mich. 2012) (applying *Daubert* to expert testimony on false confessions). *See generally*, Brian Cutler, Keith A. Findley & Danielle Loney, *Expert Testimony on Interrogation and False Confession*, 82 UMKC L. REV. 589, 590 (2014) (“The courts’ response to expert testimony on false confessions ... has not been uniformly welcoming. Some courts have permitted such evidence, but a significant number have rejected it for various reasons.”).
- 146 Primus, *supra* note 127, at 43.
- 147 David L. Faigman, John Monahan & Christopher Slobogin, *Group to Individual (G2i) Inference in Scientific Expert Testimony*, 81 U. CHI. L. REV. 417, 421 (2014) (defining framework and diagnostic evidence).
- 148 *Id.* at 440.
- 149 *See* John Monahan & Laurens Walker, *Social Authority: Obtaining, Evaluating, and Establishing Social Science in Law*, 134 U. PA. L. REV. 477, 495-97 (1986).
- 150 Faigman, Monahan & Slobogin, *supra* note 147, at 441.
- 151 *Daubert v. Merrell Dow Pharm. Co.*, 509 U.S. 579, 591 (1993) (citing *United States v. Downing*, 753 F.2d 1224, 1242 (3d Cir. 1985)).
- 152 *See* SOCIAL SCIENCE IN LAW: CASES AND MATERIALS 63-64 (John Monahan & Laurens Walker eds., 8th ed. 2014).
- 153 *Id.* at 67 (“External validity refers to the extent to which the findings of the study can be generalized.”).
- 154 *See, e.g.*, Saul M. Kassin & Katherine L. Kiechel, *The Social Psychology of False Confessions: Compliance, Internalization, and Confabulation*, 7 PSYCHOL. SCI. 125, 126 (1996) (describing a study where the consequence of the infraction was a phone call from the principal investigator); Perillo & Kassin, *supra* note 34, at 333-34 (describing a study where the consequence of the infraction was loss of one credit hour); Russano et al., *supra* note 34, at 483 (describing a study where the consequence of the infraction was either to return for another session without receiving credit or to tell the participant that the professor would be informed of the failure to confess). In none of these studies were the subjects given warnings or anything equivalent.
- 155 *See* Gisli H. Gudjonsson, *The Psychology of False Confessions: A Review of the Current Evidence*, in POLICE INTERROGATIONS AND FALSE CONFESSIONS: CURRENT RESEARCH, PRACTICE, AND POLICY RECOMMENDATIONS 31, 43 (G. D. Lassiter & C. A. Meissner eds., 2010) (“Experimental research is particularly helpful in studying the conditions under which people make false confessions and allow the researcher to control for ground truth, but this kind of research has little ecological validity in terms of applying it to real-life individual cases.”).
- 156 *See* Kassin & Kiechel, *supra* note 154, at 126. This study has been replicated by other researchers. *See* Robert Horselenberg, Harald Merkelbach & Sarah Josephs, *Individual Differences and False Confessions: A Conceptual Replication of Kassin and Kiechel (1996)*, 9 J. PSYCHOL. CRIME & L. 1, 3-5 (2003) (investigating the influence of confession consequences); Jessica R. Klaver, Zina Lee & V. Gordon Rose, *Effects of Personality, Interrogation Techniques and Plausibility in an Experimental False Confession Paradigm*, 13 LEGAL & CRIM. PSYCHOL. 71, 75-76 (2008) (investigating the effect of minimization and maximization techniques).

- 157 Kassir & Kiechel, *supra* note 154, at 127 tbl.1; *see also* Tim Cole et al., *Trying to Obtain False Confessions Through the Use of False Evidence: A Replication of Kassir and Kiechel's Study* (May 25, 2009), http://www.allacademic.com/meta/p13137_index.html [<http://perma.cc/Z5X9-RM4V>] (obtaining *no* false confessions when subjects knew they had not caused the crash); Klaver et al., *supra* note 156, at 81-82 (finding that when the “crash key” was farther away from the other keys participants were sixteen times less likely to falsely confess).
- 158 *See, e.g.*, Perillo & Kassir, *supra* note 34, at 334 tbl.2 (using the cheating paradigm and finding a significant increase in false confessions from use of bluffing technique); Russano et al., *supra* note 34, at 482, 484 tbl.1 (stating that “the participants clearly knew whether they committed the act” yet finding a significant increase in false confessions).
- 159 Russano et al., *supra* note 34, at 484 tbl.1.
- 160 Perillo & Kassir, *supra* note 34, at 334.
- 161 *See* Gudjonsson, *supra* note 155, at 43 (“In many anecdotal case studies, ground truth is difficult to ascertain. Similarly, in studies of false confessions among prisoners and community samples, the genuineness of the [self-reported] false confession is nearly impossible to corroborate.”).
- 162 *See, e.g.*, Drizin & Leo, *supra* note 5, at 929-30 (reporting 125 cases purportedly involving confessions proven false through DNA analysis or other methods); Brandon L. Garrett, *The Substance of False Confessions*, 62 STAN. L. REV. 1051, 1052-54 (2010) (reporting forty cases purporting to involve confessions that were proven false by DNA analysis).
- 163 *See* Rosenthal, *supra* note 74, at 617-18 (“[I]t would not surprise me if the vast majority of custodial interrogations involve the features condemned by critics. If so, the fact that a study of false confessions will frequently disclose the use of [manipulative] interrogation tactics ... provides no basis to conclude that these features increase the likelihood that a confession is false.”).
- 164 GARRETT, *supra* note 127, at 38.
- 165 Based on his comprehensive study, Leo concludes that these are the two most likely causes of false confessions. He asserts that negotiation is a primary cause of what the literature calls “compliant false confessions,” where the suspect confesses and subsequently recants. LEO, *supra* note 11, at 201 (describing a compliant false confession as given “to achieve some instrumental benefit--typically either to terminate and thus escape from aversive interrogation process, to take advantage of a perceived suggestion or promise of leniency, or to avoid an anticipated harsh punishment”). He concludes that the false evidence ploy, combined with prolonged interrogation, is the most common cause of “persuaded false confessions,” the second most significant category of false confessions, in which the suspect comes to believe he or she committed the crime. *Id.* at 224-25 (noting the connection between persuaded false confessions and use of the “false-evidence ploy” together with “lengthy and intense interrogation”).
- 166 Danielle E. Chojnacki, Michael D. Cicchini & Lawrence T. White, *An Empirical Basis for the Admission of Expert Testimony on False Confessions*, 40 ARIZ. ST. L.J. 1, 39 (2008) (“[O]ur survey findings indicate that the false confession phenomenon itself, even its broadest sense, is in fact *outside* the common knowledge of potential jurors.”).
- 167 Russano et al., *supra* note 34, at 484 tbl.1.
- 168 Estimates of the incidence of false confessions vary widely. *Compare* Miller W. Shealy, Jr., *The Hunting of Man: Lies, Damn Lies, and Police Interrogations*, 4 U. MIAMI RACE & SOC. JUST. L. REV. 21, 65 (2014) (“[T]he number is most likely infinitesimally small.”), *with* Kassir et al., *supra* note 139, at 5 (describing studies that found rates up to twelve percent based on self-report methods). In evaluating whether the burden of production has been met, much may depend on how courts define that threshold. *See, e.g.*, *United States v. Branch*, 91 F.3d 699, 712 (5th Cir. 1996) (“[T]he ‘merest scintilla of evidence’ in the defendant's favor does not warrant a jury instruction regarding an affirmative defense for which the defendant bears the initial burden of production [T]here must be ‘evidence sufficient for a reasonable jury to find in [the defendant's] favor.’” (quoting *United States v. Jackson*, 726 F.2d 1466, 1468 (9th Cir. 1984); then quoting *Mathews v. United States*, 485 U.S. 58, 63 (1988))).
- 169 Faigman, Monahan & Slobogin, *supra* note 147, at 451 (“While a diagnostic opinion can certainly be informed by research and the confidence level associated with it can sometimes even be quantified, ultimately whether it is reliable ... can be tested only through some sort of feedback loop that indicates whether the expert was right or wrong.”).

- 170 *Id.* at 452, 456, 464-66 (discussing how the “process of accumulating and analyzing the relevant information” can address validity concerns about diagnostic testimony).
- 171 *See, e.g.*, NAOMI E.S. GOLDSTEIN, HEATHER ZELLE & THOMAS GRISSO, *MIRANDA RIGHTS COMPREHENSION INSTRUMENTS* (2012); GISLI GUDJONSSON, *GUDJONSSON SUGGESTIBILITY SCALES* (1997); RICHARD ROGERS ET AL., *STANDARDIZED ASSESSMENT OF MIRANDA ABILITIES* (2011).
- 172 Richard A. Leo & Richard J. Ofshe, *The Consequences of False Confessions: Deprivations of Liberty and Miscarriages of Justice in the Age of Psychological Interrogation*, 88 J. CRIM. L. & CRIMINOLOGY 429, 438-39 (1998). Welsh White disagreed with this proposal on several grounds. Welsh S. White, *What Is an Involuntary Confession Now?*, 50 RUTGERS L. REV. 2001, 2024-28 (1998). But his concerns were ably rebutted by Leo and his coauthors. Leo et al., *supra* note 131, at 522-25.
- 173 Garrett, *supra* note 162, at 1066 (noting that in thirty-six of thirty-eight false confession cases suspects' confessions were “contaminated” by information fed by the police or media accounts).
- 174 *Compare* Rosenthal, *supra* note 74, at 610 (“In my experience, it was difficult to get even highly motivated cooperating defendants to remember the details of crimes they had committed.”), *with* Garrett, *supra* note 162, at 1111 (noting that in most exoneree case studies “there was a lack of fit and non-volunteered details were inconsistent with crime scene evidence”).
- 175 Garrett, *supra* note 162, at 1116 (stating that police can not only examine “whether the suspect volunteers key crime scene facts, but also [ask] leading questions regarding facts inconsistent with how the crime occurred”). Paul Cassell has rightly pointed out that many guilty suspects often provide statements, intentionally or not, that do not fit the known facts. Paul G. Cassell, *The Guilty and the “Innocent”: An Examination of Alleged Cases of Wrongful Conviction from False Confessions*, 22 HARV. J.L. & PUB. POL'Y 523, 594-96 (1999). Such a discrepancy should not be grounds for finding the confession unreliable. Rather, as indicated in (3) above, if there is no such discrepancy or only a minor one, the prosecution's burden will probably be met; if discrepancies are significant, then the court might require that (1) or (2) also be met.
- 176 Michael R. Napier & Susan H. Adams, *Criminal Confessions: Overcoming the Challenges*, 72 FBI L. ENFORCEMENT BULL. 9, 12 (2002) (noting that police routinely designate “holdback evidence” involving unique crime facts or details not publicly known or easily guessed, to see if the confessor knows about it and thereby corroborate the confession's reliability).
- 177 Lawrence Rosenthal points to George Thomas's review of custodial interrogation cases, *see* George C. Thomas III, *Stories About Miranda*, 102 MICH. L. REV. 1959, 1982-83 (2004), as evidence that admissibility decisions rarely turn on the credibility of the participants. Rosenthal, *supra* note 74, at 607. But, in fact, credibility assessments permeated Thomas's cases. *See* Thomas, *supra*, at 1975 (“[M]y data contain a potential reporting bias. If the defendant and the police tell a different story about whether warnings were given, one would expect judges to believe the police--and this is exactly what I found.”); *see also* Joseph D. Grano, *Voluntariness, Free Will, and the Law of Confessions*, 65 VA. L. REV. 859, 898 n.192 (1979) (“In most confession cases that have reached the Supreme Court, the actual events in the interrogation room have been disputed.”).
- 178 The account below captures the point:
[R]ecording will greatly facilitate the Miranda and voluntariness analyses, and a recording details factors relevant to credibility and the ultimate issue--the substance of the defendant's statements; was the defendant informed of his Miranda rights; did he understand them; were they waived; was the waiver voluntary; was the statement voluntary; was either the statement or waiver coerced; the substantive questions asked; how they were asked; and conversely the answers given and how the responses were made; the interrogator's demeanor (and appearance) contrasted with the suspect's behavior (and appearance); the fit between what the tape reveals and the testimony of the people on the tape; as the Eighth Circuit recognized, a tape will display if the defendant “is hesitant, uncertain, or faltering, ... [if] he has been worn out by interrogation, physically abused, or in other respects is acting involuntarily, the tape will corroborate him in ways a typewritten statement would not.
Daniel Donovan & John Rhodes, *Comes a Time: The Case for Recording Interrogations*, 61 MONT. L. REV. 223, 228 (2000).
- 179 FED. R. CIV. P. 30(c).
- 180 Slobogin, *supra* note 9, at 317.
- 181 *False Confessions & Recording of Custodial Interrogations*, INNOCENCE PROJECT (Aug. 12, 2015), www.innocenceproject.org/free-innocent/improve-the-law/fact-sheets/false-confessions-recording-of-custodial-

interrogations [<https://perma.cc/3LMF-9X8C>] (reporting that approximately one thousand jurisdictions in the United States use interrogation room recording equipment); Carrie Johnson, *New DOJ Policy Urges Agents to Videotape Interrogations*, NAT'L PUB. RADIO (May 21, 2014, 5:04 PM), <http://www.npr.org/blogs/thetwo-way/2014/05/21/314616254/new-doj-policy-calls-for-videotaping-the-questioning-of-suspect> [<https://perma.cc/8EMJ-Q4VU>] (noting a reversal of federal policy against recording interrogations).

- 182 Brandon L. Garrett, *Interrogation Policies*, 49 U. RICH. L. REV. 895, 898 (2014) (finding that only eight percent of 116 Virginia departments required recording); Andrew E. Taslitz, *High Expectations and Some Wounded Hopes: The Policy and Politics of a Uniform Statute on Videotaping Custodial Interrogations*, 7 NW. J.L. & SOC. POL'Y 400, 409 (2012) (“[T]he vast majority of police departments still do not record. There are wide variations among the voluntarily adopted programs. Departments vary in what crimes are recorded, whether recording is only audio or also visual, and at what locations recording must be made.”); Tracy Lamar Wright, *Let's Take Another Look at That: False Confession, Interrogation, and the Case for Electronic Recording*, 44 IDAHO L. REV. 251, 279, 281 (2007) (“[M]any of the nation's largest police departments that do record interrogations only record the latter part where the suspect confesses” and many record only interrogations in homicide or serious felony cases.”).
- 183 See Slobogin, *supra* note 9, at 317-21.
- 184 At least one court has accepted this argument. *Stephan v. State*, 711 P.2d 1156, 1162 (Alaska 1985). Other courts have more or less followed suit, albeit bottoming their conclusion on their supervisory power rather than the Constitution. See, e.g., *Commonwealth v. DiGiambattista*, 813 N.E.2d 516, 533 (Mass. 2004) (“[A] defendant whose interrogation has not been reliably preserved by means of a complete electronic recording should be entitled, on request, to a cautionary instruction concerning the use of such evidence.”); *State v. Barnett*, 789 A.2d 629, 632 (N.H. 2001) (“[A] tape recorded interrogation will not be admitted into evidence unless the statement is recorded in its entirety.”).
- 185 467 U.S. 479 (1984).
- 186 488 U.S. 51 (1988).
- 187 *Id.* at 58; *Trombetta*, 467 U.S. at 490-91.
- 188 See, e.g., *Holloway v. Horn*, 161 F. Supp. 2d 452, 530 (E.D. Pa. 2001) (stating that a recording of an interrogation is only “potentially useful”); *Tennessee v. Godsey*, 60 S.W.3d 759, 771 (Tenn. 2001) (“Lack of an electronic recording did not preclude the defendant from challenging the accuracy of the officers' recollection of the interrogation.”).
- 189 John H. Langbein, *The Historical Origins of the Privilege Against Self-Incrimination at Common Law*, 92 MICH. L. REV. 1047, 1059-60 (1994).
- 190 Roger Lane, *Urban Police and Crime in Nineteenth Century American*, in 15 MODERN POLICING 1, 5 (Michael Tonry & Norval Morris eds., 1992) (explaining that in the nineteenth century, law enforcement was “largely the responsibility either of the community as a whole or of the individual victim of some offense, rather than something delegated to specialized agents of the state”).
- 191 In *Toward Taping*, I also argued, based on the holding in *Connelly*, that taping is required given the prosecution's burden of showing by a preponderance of the evidence that the Fifth Amendment was not violated: If one assumes that voluntariness cannot be assessed without taping, the tapeless prosecutor cannot meet [the preponderance of the evidence] burden, at least where the defendant plausibly asserts he did not receive or understand warnings, was misled about them, or received improper threats, promises and the like. In such cases, at best the parties are in equipoise, and the party with the burden of proof--the government--should lose. Slobogin, *supra* note 9, at 319.
- 192 388 U.S. 218 (1967).
- 193 *Id.* at 237.
- 194 *Id.* at 228.

- 195 *Id.* at 235.
- 196 *See, e.g., United States v. LaPierre*, 998 F.2d 1460, 1464 (9th Cir. 1993) (finding insufficient a videotape that showed only the lineup and not what occurred in the witness room); *People v. Fowler*, 461 P.2d 643, 654 (Cal. 1969) (finding still photographs inadequate); *Bruce v. Indiana*, 375 N.E.2d 1042, 1086 (Ind. 1978) (requiring videotaping lineups).
- 197 Slobogin, *supra* note 9, at 321.
- 198 *See* Richard A. Leo, *Inside the Interrogation Room*, 86 J. CRIM. L. & CRIMINOLOGY 266, 276 (1996) (finding a seventy-eight percent waiver rate).
- 199 *Pate v. Robinson*, 383 U.S. 375, 384-85 (1966) (requiring inquiry into competence to stand trial upon any “bona fide doubt”).
- 200 *See* Nancy J. King, *Priceless Process: Nonnegotiable Features of Criminal Litigation*, 47 UCLA L. REV. 113, 117 (1999) (“[R]ules of constitutional stature protecting interests that may differ from those of the parties should not be subject to evasion by the consent of the parties unless effective enforcement mechanisms exist to protect such interests.”).
- 201 Some predict such cameras will become as common as dashboard cameras within the very near future. *See* Chuck Humes, *Body Worn Cameras*, LAW OFFICER (Feb. 14, 2013), [www.lawofficer.com/article/technology-and-communications/body-worn-cameras](https://perma.cc/P4Q3-JJHR) [https://perma.cc/P4Q3-JJHR]. Even expense might not be a major limitation. *See Justice Department Awards over \$23 Million in Funding for Body Worn Camera Pilot Program to Support Law Enforcement Agencies in 32 States*, DEPT JUST. (Sept. 21, 2015), <https://www.justice.gov/opa/pr/justice-department-awards-over-23-million-funding-body-worn-camera-pilot-program-support-law> [https://perma.cc/V6W4-B85Q] (describing the Department of Justice initiative “to assist local jurisdictions that are interested in exploring and expanding the use of body-worn cameras”).
- 202 *See generally* Barak Ariel et al., *The Effect of Police Body-Worn Cameras on Use of Force and Citizens' Complaints Against the Police: A Randomized Controlled Trial*, 31 J. QUANTITATIVE CRIMINOLOGY 509, 525-26 (2015) (finding that use-of-force complaints against the police who used body cameras was roughly half the number lodged against police without cameras, perhaps because both police and citizens were more circumspect).
- 203 *Oregon v. Elstad*, 470 U.S. 298, 314 (1985) (holding that although a postwarning statement that repeats a prewarning statement is usually admissible, it is inadmissible if the police use “deliberately coercive or improper tactics in obtaining the initial statement”); *see also, Missouri v. Seibert*, 542 U.S. 600, 621 (2004) (Kennedy, J., concurring) (while affirming *Elstad*, stating that when an interrogator deliberately uses the “two-step strategy, predicated upon violating *Miranda* during an extended interview, postwarning statements that are related to the substance of prewarning statements must be excluded absent specific curative steps”); *United States v. Patane*, 542 U.S. 630, 632 (2004) (where five justices indicated bad faith might require exclusion even of tangible fruits).
- 204 There may be good reasons to be careful about adopting such a system. *See* Elizabeth Atkins, *#BlackLivesRecorded: Will the Darling Savior of Police Brutality Be the Downfall of Modern Privacy?* 13-14 (2016), https://papers.ssrn.com/sol3/papers2.cfm?abstract_id=2803588 [https://perma.cc/CX8Q-Y57Q] (delineating privacy and other harms that can arise from body camera use). While the argument in the text does not require adoption of such a system, it would prohibit police questioning without it.
- 205 U.S. CONST. amend. VI.
- 206 *See, e.g., Lilly v. Virginia*, 427 U.S. 116, 117-18 (1999) (holding a codefendant's confession made before the defendant was inadmissible under the Confrontation Clause).
- 207 U.S. CONST. amend. V.
- 208 *Miranda v. Arizona*, 384 U.S. 436, 444 (1966) (stating that the Fifth Amendment applies to “custodial interrogation,” defined as “questioning initiated by law enforcement officers after a person has been taken into custody or otherwise deprived of his freedom of action in any significant way”).
- 209 *Arizona v. Fulminante*, 499 U.S. 279, 286-87 (1991) (applying due process analysis to interaction that did not involve custody).

- 210 See, e.g., Norman Abrams, *The Case for a Cabined Exception to Coerced Confession Doctrine in Civilian Terrorism Prosecutions*, in PATRIOTS DEBATE: CONTEMPORARY ISSUES IN NATIONAL SECURITY LAW 42, 51 (Harvey Rishikof, Stewart Baker & Bernard Horowitz eds., 2012) (arguing that some interrogation techniques that are considered coercive under current doctrine should be permissible in the national security context under an “exigent circumstance exception”).
- 211 See *Chavez v. Martinez*, 538 U.S. 760, 776 (2003).
- 212 See *supra* text accompanying notes 23; see also Bobby Ghosh, *After Waterboarding: How to Make Terrorists Talk?*, TIME (June 8, 2009), www.time.com/time/magazine/article/0,9171,1901491,00.html [<https://perma.cc/QM5U-E9WM>] (stating that, according to government interrogators, “the best way to get intelligence from even the most recalcitrant subject is to apply the subtle arts of interrogation,” including treating the individual with respect and then using “sleight of hand” to get the relevant information).
- 213 467 U.S. 64, 655-56 (1984) (announcing a “‘public safety’ exception” to *Miranda*).
- 214 Abrams, *supra* note 210, at 49-50; see also *F.B.I. Memorandum*, N.Y. TIMES (Mar. 25, 2011), www.nytimes.com/2011/03/25/us/25miranda-text.html?_r=1 [<https://perma.cc/QR9B-4W8G>] (providing the text of an internal Justice Department memorandum that permits agents to decide whether “continued unwarned interrogation is necessary to collect valuable and timely intelligence not related to any immediate threat”).
- 215 *Quarles*, 467 U.S. at 658 (emphasizing that the officers in the case “were confronted with the immediate necessity of ascertaining the whereabouts of a gun which they had every reason to believe the suspect had just removed from his empty holster and discarded in the supermarket”).
- 216 *Id.* at 655 n.5 (“[R]espondent is certainly free on remand to argue that his statement was coerced under traditional due process standards.”).
- 217 *Hamdi v. Rumsfeld*, 542 U.S. 507, 533-35 (2004) (holding that the “exigencies” of a military trial allow departure from normal procedures other than the “core elements” of “notice ... and a fair opportunity to rebut the Government’s factual assertions before a neutral decisionmaker”).
- 218 See Arnold H. Loewy, *Police Obtained Evidence and the Constitution: Distinguishing Unconstitutionally Obtained Evidence from Unconstitutionally Used Evidence*, 87 MICH. L. REV. 907, 939 (1989) (“[W]hen obtaining evidence is the constitutional wrong [as opposed to when the wrong is using it in a criminal proceeding], exclusion should be subjected to a cost-benefit analysis.”).
- 219 *Id.* at 938-39 (arguing that deterrence is the main goal of the Due Process Clause and concluding that, given that goal, third party standing might be granted “to deter the most flagrant forms of obtaining coerced confessions”). However, a defense might be available in such situations. In Israel, interrogators can resort to any means needed to procure information that might avert a threat, subject to the stipulation that they will escape subsequent prosecution or suit only if they can prove a necessity defense. Judgment of the Interrogation Methods Employed by the General Security Service, Israeli Supreme Court ¶¶ 35-36 (1999), <https://www.derechos.org/human-rights/mena/doc/torture.html> [<https://perma.cc/9SAY-VWDC>].

97 BULR 1157

End of Document

© 2017 Thomson Reuters. No claim to original U.S. Government Works.

Traces of Crime: How New York's DNA Techniques Became Tainted

The city's medical examiner has been a pioneer in analyzing complex DNA samples.

But two methods were recently discontinued, raising questions about thousands of cases.

By LAUREN KIRCHNER SEPT. 4, 2017

Over the past decade, the DNA laboratory in the office of New York City's chief medical examiner emerged as a pioneer in analyzing the most complicated evidence from crime scenes. It developed two techniques, which went beyond standard practice at the F.B.I. and other public labs, for making identifications from DNA samples that were tiny or that contained a mix of more than one person's genetic material.

As its reputation spread, the lab processed DNA evidence supplied not only by the New York police, but also by about 50 jurisdictions as far away as Bozeman, Mont., and Floresville, Tex., which paid the lab \$1,100 per sample.

Now these DNA analysis methods are under the microscope, with scientists questioning their validity. In court testimony, a former lab official said she was fired for criticizing one method, and a former member of the New York State Commission on Forensic Science said he had been wrong when he approved their use. The first expert witness allowed by a judge to examine the software source code behind one technique recently concluded that its accuracy "should be seriously questioned."

9

ARTICLES REMAINING

[SEE MY OPTIONS](#)[Subscriber login](#)

A coalition of defense lawyers is asking the New York State inspector general's office — the designated watchdog for the state's crime labs — to launch an inquiry into the use of the disputed analysis methods in thousands of criminal cases. While the inspector general has no jurisdiction over the court system, any finding of flaws with the DNA analysis could prompt an avalanche of litigation. Previous convictions could be revisited if the flawed evidence can be shown to have made a difference in the outcome.

The medical examiner's office "has engaged in negligent conduct that undermines the integrity of its forensic DNA testing and analysis," the Legal Aid Society and the Federal Defenders of New York wrote the inspector general on Friday. Because the lab has kept problems with its "unreliable" testing and "unsound statistical evidence" secret from the public and the courts, they continued, "innocent people may be wrongly convicted, and people guilty of serious crimes may go free."

In addition to those convicted using the disputed methods, many defendants may have chosen to plead guilty when they learned prosecutors had DNA evidence against them. Their cases face significant barriers to reconsideration.

The medical examiner's office stands by its science. Its chief of laboratories, Timothy Kupferschmid, said that the discarded techniques were well-tested and valid, and that the lab was adopting newer methods to align with changing F.B.I. standards. He compared it to a vehicle upgrade.

"So just because we're switching to the new model, I mean, our old pickup truck worked great, but my new pickup truck is so much better," he said.

One case that hinges on the disputed DNA techniques stemmed from the beating of Taj Patterson in December 2013. A group of Hasidic men attacked Mr. Patterson, a black student, in the Williamsburg section of Brooklyn. Prosecutors blamed the attack on the Shomrim, a Hasidic group that patrols Williamsburg, a neighborhood where tensions between orthodox Jews and blacks have long simmered.

Six days after the attack, the police found one of Mr. Patterson's black Air Jordan sneakers on a nearby roof.

The police sent the sneaker to the DNA lab, where a technician swabbed a 3-inch by 6-inch area of its heel — and recovered 97.9 picograms of DNA from at least two people. A picogram is one trillionth of a gram.

The sample bore Mr. Patterson's DNA. Using software developed in-house, the lab calculated that it was 133 times more likely than not that the remainder belonged to Mayer Herskovic, a young father who lived and worked in Williamsburg and had no criminal record.

"I don't believe that this is DNA," Mr. Herskovic said. "A mixture, like you take milk, orange juice and water and you mix it, what is it? Is it still milk? Is it still orange juice? I don't know."

"DNA is the magic word," he added. "If you throw it into a trial, they eat it up. For me, it's not magic at all."

No other physical evidence linked Mr. Herskovic to the attack on Mr. Patterson, who was blinded in his right eye. Neither the victim nor those who witnessed the crime identified Mr. Herskovic at trial, nor was he seen on surveillance video. Mr. Herskovic said he has never been part of the Shomrim, and deplored the assault on Mr. Patterson.

Nevertheless, he was convicted by a judge of gang assault, and sentenced this past March to four years in prison. He is appealing.

Three years ago, Barry Scheck, a co-founder of the Innocence Project, a nonprofit that uses DNA evidence to exonerate wrongly convicted prisoners, yelled at his colleagues on the state forensic commission about the potential perils of the DNA work at the city's lab.

"The day of reckoning is going to come," Mr. Scheck told his fellow commissioners, some of whom rolled their eyes, a video of the meeting showed. "Someday people are going to review this," he continued. "It's an Ebola. It is a cancer here that could be spreading. We are all on notice."

• • •

For three decades, forensic DNA evidence has been a valuable tool in criminal investigations, incriminating or exonerating suspects. Matching a defendant's genetic material with a sample found on a weapon or at a crime scene has proved extremely persuasive with judges and juries.

But not all DNA evidence is equal. Sometimes it's clear: blood or semen identifies a single person. If it's just a few skin cells left on an object, or if it contains more than one person's genetic material, it can be more ambiguous. In such situations, labs used to report that the results were inconclusive, or the defendant could not be excluded from the mix.

New types of DNA analysis have been introduced in recent years to interpret trace amounts or complex mixtures, spawning an industry of testing tools, chemical kits and software. As analysis has become more complex, the techniques and results are coming under fire nationwide.

In the past three years, flaws in DNA methods have temporarily shut down testing in public crime labs in Austin, Tex., and Washington. Lab analysts "make it seem like it's a completely objective process," said Bicka Barlow, a lawyer in California with a master's degree in genetics and molecular biology. "But I'm 100 percent convinced that there are many people who are incarcerated who were convicted with DNA evidence who are innocent."

The two techniques that New York's lab introduced were the "high-sensitivity testing" of trace DNA amounts, and the Forensic Statistical Tool, or FST, in which software calculates the likelihood that a suspect's genetic material is present in a complicated mixture of several people's DNA. By its own estimate, the lab has used high-sensitivity DNA testing to analyze evidence samples in 3,450 cases over the past 11 years, and the FST in 1,350 cases over the past six. Cases in which both methods were used may be counted in both totals.

• • •

In February 2012, responding to a 911 call about gunshots near East Tremont Avenue, police officers from the 45th Precinct in the Bronx saw a passer-by make a motion as if he was dropping an object under a parked car.

His was a familiar face: **Johnny Morgan**, who had been arrested 75 times. The police found a .40-caliber Glock 23 beneath the car. Mr. Morgan was charged with gun possession, based both on DNA evidence and witness testimony. But the amount of DNA recovered from the gun was extremely small; the lab initially said it was unsuitable for testing.

After the prosecutor and the police requested a high-sensitivity test, analysts said Mr. Morgan's DNA was a match. He was convicted.

Public crime labs assessing DNA evidence, including the F.B.I.'s lab, "amplify," or copy, the material 28 times to conduct their analysis. Under the high-sensitivity testing method developed by Dr. Theresa A. Caragine, a forensic scientist, and implemented in 2006, New York's lab began to push very small amounts through three more cycles, bringing the total to 31. This approach provided more material to look at — as much as eight times the standard approach. But, like turning up the volume on a radio, those additional cycles amplified small imperfections from missing or contaminated DNA.

To reduce potential problems, the lab decided not to amplify samples smaller than 20 picograms, or about three cells' worth of DNA, its then-director, Dr. Mechthild Prinz, said in 2005 during the state's approval process for the test. She declined to comment for this article.

"The scientific community has been asked to test more and more evidence with less and less amounts of DNA," Dr. Prinz explained in 2009 to the DNA Subcommittee of the state forensic science commission, which approves all forensic methods used in New York State.

"A couple of years ago, DNA testing was limited to body fluids — semen, blood, and saliva. Now every laboratory in the country routinely receives swabs from guns," other weapons, burglary tools, and cash registers, she said.

After several years of high-sensitivity testing of small amounts of DNA, the lab developed a second method: a piece of software to interpret complex mixtures.

Invented by Dr. Caragine and Dr. Adele A. Mitchell, a geneticist with a specialty in statistics who joined the lab in 2008, the Forensic Statistical Tool, or FST, considers the overall amount of DNA in the mixture, how many people are in it, how much information is probably missing or contaminated, and the frequency with which each piece of DNA appears in different racial or ethnic groups. Then it compares the defendant's DNA profile to the mixture, and calculates a likelihood ratio, which it expresses as a single number.

The bigger that number — and it's sometimes in the millions or even trillions — the more likely that the defendant's DNA is present. Dr. Caragine and Dr. Mitchell testified in 2012 that about a third of all test results were favorable to defendants, by indicating that their DNA was probably absent.

Only a small proportion of cases using the Forensic Statistical Tool went to trial. Most defendants faced with unfavorable FST results pleaded guilty, defense lawyers say. “Just the prospect of those numbers going in front of the jury could really warp the plea bargaining process,” said Brad Maurer, a lawyer and DNA specialist at New York County Defender Services.

Eric Rosenbaum, an assistant district attorney and head of the DNA Prosecutions Unit in Queens, described the FST as an “extremely powerful tool because it is devastating in court.”

• • •

In December 2012, The New York Times profiled Dr. Mitchell and Dr. Caragine in the article “Helping Decide Guilt or Innocence,” which described their fruitful collaboration, but also hinted at a brewing controversy. The Legal Aid Society was gearing up for an extensive fight against admission of FST results in court.

One interested reader was Dr. Eli Shapiro, the former mitochondrial DNA technical leader in the DNA lab. One reason for his early retirement, he later testified, was the stress over having to sign off on lab reports generated by the software. Even in the lab, few people knew the science behind it.

Dr. Shapiro later said in court that he found the FST process described in the article “very disturbing.” He reached out to his former boss and colleagues to express his alarm. “They were not concerned,” he testified.

So, in early 2013, Dr. Shapiro offered his help to Legal Aid, which had just formed a unit specializing in DNA evidence. Under a judge’s order, the lab had given Legal Aid the results of its validation studies — internal tests of the FST’s accuracy. Dr. Shapiro helped decipher the data.

“He knows the math,” said Clinton Hughes, a Legal Aid lawyer. “For relaxation, he does long division on the beach with a pencil.”

From 2012 to 2014, a hearing in Brooklyn before Judge Mark Dwyer focused on DNA evidence in two cases: it had been recovered from the handlebars of a bicycle after a shooting, and from the clothing of a sexual assault victim. With the help of testimony from Dr. Shapiro and some of the world’s most renowned DNA experts, Legal Aid hoped to persuade the judge to throw out the evidence.

The defense experts were denied access to the FST’s software code, which would later come under scrutiny. Instead, they criticized the way that Dr. Caragine and Dr. Mitchell designed and tested the FST.

Dr. Bruce Budowle, an architect of the F.B.I.’s national DNA database, testified that New York’s statistical methods were “not defensible.”

He said that the FST was designed with the incorrect assumption that every DNA mixture of the same size was missing information or had been contaminated in just the same way. He also criticized the lab’s overreliance on “pristine” saliva and samples to test its methods, which do not mirror the ways real crime-scene evidence is degraded by time and weather. The lab underestimated the challenges, he testified.

“Five-person mixtures can look like three-person,” he said, “four contributors can look like two-person mixtures. It’s almost impossible to actually be accurate.”

The software’s inventors acknowledged a margin of error of 30 percent in their method of quantifying the amount of DNA in a sample, a key input into the FST

calculation. They acknowledged that the FST didn't consider that different people in a mixture, especially family members, might share DNA.

In April 2013, weeks after testifying, Dr. Caragine was forced to resign from the lab after New York's inspector general found that she had violated protocol by changing her colleagues' FST results in two cases. Her defense was that she was correcting their mistakes. Dr. Mitchell left in 2014. Dr. Caragine declined to comment for this article, and Dr. Mitchell did not respond to repeated requests for comment.

Perhaps the most dramatic testimony in the hearing came from Dr. Ranajit Chakraborty, who had developed the F.B.I.'s policy on DNA in the 1990s and, as a member of New York's DNA Subcommittee, voted to approve both high-sensitivity testing in 2005 and the FST in 2010. What he had since learned about the FST bothered him.

"What would your vote be today?" Jessica Goldthwaite, a lawyer for Legal Aid, asked Dr. Chakraborty on the stand.

"My answer would be no," he said. In November 2014, Judge Dwyer sided with the defense, excluding evidence produced by both high-sensitivity testing and the FST. He was the first state judge to do so, and so far the only one.

• • •

Appointed to the state forensic science commission when it formed in 1994, Mr. Scheck didn't vote for either of the lab's methods. His misgivings grew when he learned that the DNA sample used to convict Mr. Morgan in the Bronx gun case was only 14.15 picograms. That was below the 20-picogram minimum for high-sensitivity testing the lab had promised to set during its approval process back in 2005.

At the October 2014 commission meeting, Mr. Scheck pounded the table as he proposed to compel the lab to turn over any validation studies it had conducted for high-sensitivity testing of especially small samples. He accused lab officials of not having performed the necessary studies, despite their assurances otherwise. While Mr. Scheck's motion failed, it drew a vote from an unexpected supporter: Dr. Marina

Stajic, who then worked for the medical examiner's office as the director of the toxicology lab. She supported the motion, she later testified, because she believed that the DNA lab should be transparent with its data.

Her boss, Dr. Barbara Sampson, the chief medical examiner, heard about Dr. Stajic's vote the next morning. She expressed her anger in an email to a colleague, "Hold me down."

Mimi Mairs, then a lawyer for the DNA lab, emailed, "She sucks."

A spokeswoman for the medical examiner's office declined to comment on the correspondence, as did the Manhattan district attorney's office, where Ms. Mairs is now a prosecutor.

In April 2015, Dr. Sampson and Mr. Kupferschmid fired Dr. Stajic, who had worked at the lab for 29 years. Mr. Kupferschmid then called a commission member to inquire whether Dr. Stajic would also be removed from the oversight group, according to court documents.

In February 2016, Dr. Stajic sued Dr. Sampson, Mr. Kupferschmid, and the city for allegedly violating her First Amendment rights. The defendants' lawyer contends Dr. Stajic can't prove why she was fired, and that her vote wasn't constitutionally protected speech. Her case is pending.

• • •

The case that finally revealed the FST's source code began with a few drops of cooking oil.

Kevin Johnson and his ex-girlfriend Octavia Martin shared custody of two sons, and he sometimes stayed over in her Bronx apartment. One night in April 2015, he was cooking cheeseburgers when some oil spilled. He and Ms. Martin argued about cleaning it up.

Her daughter got upset and called 911, telling the dispatcher that Mr. Johnson was pointing a gun at Ms. Martin. A police search of the apartment turned up two

socks wedged between the refrigerator and the wall. In one sock was a black pistol; in the other, a silver revolver.

Mr. Johnson, who had been convicted on a previous weapons charge, was arrested.

The lab found that one gun contained two people's DNA; by the FST's calculation, it was 156 times more likely than not to contain Mr. Johnson's DNA. The second gun had three people's DNA and a formidable likelihood of 66 million.

Hoping to cast doubt on the DNA results, his lawyers, Christopher Flood and Sylvie Levine, asked for the FST source code, which other lawyers had sought in vain.

Again, the government refused to hand it over on the grounds that it was a "proprietary and copyrighted" statistical tool owned by the City of New York.

The federal judge granted the defense access to the FST code in June 2016 under an order that bars wider disclosure. (The medical examiner's office denied ProPublica's public records request for the code, citing its "sensitive nature.")

Nathaniel Adams, a computer scientist and an engineer at a private forensics consulting firm in Ohio, reviewed the code for the defense. He found that the program dropped valuable data from its calculations, in ways that users wouldn't necessarily be aware of, but that could unpredictably affect the likelihood assigned to the defendant's DNA being in the mixture.

"I did not leave with the impression that FST was developed by an experienced software development team," Mr. Adams wrote in an affidavit. Pending more rigorous testing, "the correctness of the behavior of the FST software should be seriously questioned." Characterizing Mr. Adams' criticisms as merely stylistic rather than substantive, the lab told ProPublica that the FST provided reliable calculations.

Technology consultants wrote the software code for the FST, according to a spokeswoman at the medical examiner's office. Few, if anyone, at the lab or on the state's DNA Subcommittee had the expertise to double-check the software, said a scientist in the lab who worked on the techniques who asked to remain anonymous for fear of career repercussions. "We don't know what's going on in that black box,

and that is a legitimate question,” the scientist said, adding that evidence in older cases should “absolutely” be retested in light of growing questions about the FST. “As a scientist, I can’t say no.”

The U.S. attorney’s office withdrew the DNA evidence against Mr. Johnson days before the hearing about its admissibility was scheduled to begin.

Nevertheless, Mr. Johnson pleaded guilty this past May. On Aug. 28, he was sentenced to 28 months in prison, almost all of which he has already served. His lawyers declined to make him available for an interview.

As Mr. Johnson’s case proceeded, the lab circulated a memo to clients in September 2016, notifying them that it would replace both high-sensitivity testing and the FST on Jan. 1. A new chemical kit would make the additional amplification cycles of the high-sensitivity method unnecessary. The lab would retire the FST in favor of STRmix, a commercially available and F.B.I.-endorsed software program for DNA mixtures that dozens of public labs use.

The medical examiner’s office “is fully committed to staying on the cutting edge of new technology to best serve the City of New York,” Mr. Kupferschmid wrote in the memo. He added that the lab would raise the minimum sample size for sensitivity testing to 37.5 picograms — almost twice the initial floor of 20 picograms.

• • •

The change in policy is scant consolation to those who were convicted based on the discarded DNA techniques, like Mr. Herskovic. After the gang attack on Mr. Patterson, two confidential informants gave Mr. Herskovic’s name to a police detective. Mr. Herskovic was then arrested and swabbed for DNA. Neither informant testified against him at trial.

Sitting at a table in his apartment in Williamsburg, Mr. Herskovic discussed the DNA evidence, first calmly and then indignantly. The white walls were bare except for a small mirror, a clock and a portrait of his children, who were scribbling in coloring books on the kitchen floor. He recalled how, when the police asked him to

give a DNA sample, his lawyer cautioned him not to, but Mr. Herskovic went ahead and did so.

“I was the first one to give DNA,” Herskovic said. “He told me they needed it, I said, ‘Go ahead, take it! It will be better.’”

The DNA on Mr. Patterson’s sneaker was pivotal to the case against Mr. Herskovic. Mr. Patterson testified that whoever pulled off his shoe had punched and kicked him. Although four other suspects were arrested, and several other men were identified by witnesses, seen on surveillance video, or had their license plates photographed at the scene, only Mr. Herskovic has been tried or sentenced to prison. Two people pleaded guilty to misdemeanors and were given probation; charges were dropped against the other two.

Mr. Herskovic’s four-year sentence was stayed pending appeal. He’s working at an hourly job for a heating, ventilating and air-conditioning company to support his wife and two young children. His appeals lawyer, Donna Aldea, plans to argue that the FST was never tested on a population as insulated as the Hasidic Jews of Williamsburg, who very likely share many of the same ancestors, and therefore much of the same DNA.

“This case is a poster-child for how ‘DNA evidence’ can literally be fabricated out of thin air, and how statistics can be manipulated to create a false impression of ‘scientific evidence’ of guilt,” Ms. Aldea said. “This must be exposed.”

Lauren Kirchner is a senior reporting fellow for ProPublica, an independent, nonprofit newsroom that produces investigative journalism in the public interest.

A version of this article appears in print on September 5, 2017, on Page A1 of the New York edition with the headline: Doubts and DNA Evidence.

2007 Wis. L. Rev. 701

Wisconsin Law Review
2007

Comment
Molly Gena ^{al}

Copyright (c) 2007 By the Board of Regents of the University of Wisconsin System; Molly Gena

SYNDROME: MEDICAL UNCERTAINTY CASTS DOUBT ON CONVICTIONS

I.	Introduction	702
II.	What Is Shaken Baby Syndrome?	705
	A. Symptoms of Shaken Baby Syndrome: The Triad	707
	B. Summary of Significant SBS Convictions	708
	1. Earliest Cases	708
	2. Commonwealth v. Woodward	709
	C. The Changing Science	710
	1. Other Causes of SBS Symptoms	710
	2. The Existence of Lucid Intervals	712
III.	The Role of Medical Experts in SBS Trials	714
	A. Legal Standard Beyond a Reasonable Doubt	714
	B. Reasonable Medical Certainty	715
	C. Role of Doctors in SBS Cases	718
IV.	Looking to Other Countries for Solutions	720
V.	Conclusion	725

***702** If murder cannot be proved, the conviction cannot be safe. In a criminal case, it is simply not enough to be able to establish even a high probability of guilt. Unless we are sure of guilt the dreadful possibility always remains that a mother, already brutally scarred by the unexplained death or deaths of her babies, may find herself in prison for life for killing them when she should not be there at all. In our community, and in any civilised community, that is abhorrent. ¹

I. Introduction

On April 3, 2004, emergency personnel received a call that six-month-old Riley Owen Bilke was “red in the face and breathing heavily.” ² Later that day, despite rescue efforts, he died from symptoms commonly associated with **shaken baby** syndrome (SBS). ³ His father, Todd Dittberner, was charged with first-degree reckless homicide. ⁴ The state's medical experts testified that Riley suffered from the triad of symptoms that traditionally lead to a diagnosis of SBS-- brain hemorrhaging, retinal hemorrhaging, and brain swelling. ⁵ They suggested, as is typically assumed with SBS, ⁶

SHAKEN BABY SYNDROME: MEDICAL UNCERTAINTY..., 2007 Wis. L. Rev. 701

that the baby was probably crying inconsolably, and that Dittberner became so frustrated that he **shook** the **baby** to death.⁷ They concluded that only SBS could have caused those symptoms.⁸

Dittberner, however, maintained that he did not shake his son.⁹ He said that, when he was burping his son, the child had a seizure,¹⁰ “stiffened up,” and slipped out of his arms.¹¹ He asserted that the child *703 fell on the floor and landed on the top of his head.¹² Dittberner was alone with the baby, except for a two-year-old child he was babysitting at the time.¹³ Upstairs neighbors--who claimed that they could hear everything in Dittberner's apartment--stated that they heard neither the baby crying nor any yelling on the day of Riley's death.¹⁴

Deborah Crawford was one of the first emergency workers to arrive on the scene.¹⁵ Testifying for the prosecution, she stated that Riley's mother was upset, but that Dittberner was “emotionless.”¹⁶ She did not think that the baby could have sustained such injuries from a short fall, noting that “generally a baby doesn't fall such a small amount and go unconscious.”¹⁷

In his defense, Dittberner presented several experts who stated that the type of injuries that Riley exhibited could have been caused by a combination of other factors.¹⁸ Horace Gardner, “one of the country's foremost experts on the human eye,”¹⁹ testified that the baby would not have sustained the type of injuries that he found if the child had been shaken to death.²⁰ Although Riley had retinal hemorrhaging,²¹ Gardner explained that retinal injuries more commonly result from a lack of oxygen.²² He testified that he had not seen any cases of SBS in which the victim exhibited the particular type of retinal hemorrhages that Riley had,²³ and suggested that the injury could have come from an improper insertion of the breathing tube during the emergency treatment.²⁴ Apparently, the emergency personnel had serious difficulty trying to intubate the baby, depriving him of oxygen.²⁵

The defense also introduced the testimony of Patrick Barnes, a pediatric neuroradiologist and child-abuse expert.²⁶ He testified that Riley had a history of health problems, including bronchitis and *704 pneumonia.²⁷ Barnes stressed that doctors should not accuse a parent or caregiver of child abuse until they have ruled out every other possibility.²⁸ Because the Dittberners had adopted Riley, doctors did not know all of his medical history, nor did they perform certain tests that might have helped determine Riley's cause of death.²⁹

Although Dittberner was ultimately acquitted,³⁰ his case exemplifies the difficulties with charges of child abuse based on allegations of SBS. Often these cases become a battle of experts in which the state's witnesses claim that the child must have died from being violently shaken, and the defense's (if the defense can retain them)³¹ provide alternative explanations for the symptoms.³² It may be difficult for the fact-finder when there is no corroborating **evidence**, and only medical testimony that the child suffered from the triad of SBS symptoms.³³ The jury, possibly confused due to the complicated nature of an SBS diagnosis,³⁴ might return a guilty verdict because of the nature of the crime--the tragic death or severe injury of a baby³⁵; on the other hand, the jury might acquit, wanting to believe that no one could intentionally harm an infant.³⁶ Dittberner was fortunate because his neighbors were able to testify in his favor and he had the support of some of the leading medical experts in the country.³⁷

SHAKEN BABY SYNDROME: MEDICAL UNCERTAINTY..., 2007 Wis. L. Rev. 701

SBS cases differ from other murder or injury cases both in that they are often based only on medical opinions about the triad, and that there may not have been a crime committed at all.³⁸ Thousands of Americans *705 have been sent to prison over the last two decades on charges related to SBS,³⁹ and most are convicted only on the existence of the SBS triad of symptoms and the fact that they were with the baby when it became fatally ill.⁴⁰ The usual lack of additional eyewitnesses further complicates these cases.⁴¹

This Comment explores criminal charges based on SBS and the potential for **wrongful convictions**. Part II introduces the medical terminology, gives a brief legal history of cases involving SBS, and discusses the changing science behind the syndrome. Part III outlines the scope and role of medical experts in SBS trials. In particular, it compares the term “beyond a reasonable doubt” with “reasonable medical certainty,” and determines that the commonly used terms are not compatible. Finally, Part IV looks to the examination and management of SBS cases in other countries.

This Comment concludes that infant deaths can lead to the homicide convictions of innocent people. The medical community does not agree that subdural hemorrhaging and retinal hemorrhaging without **evidence** of an impact are necessarily indicative of shaking.⁴² The standard of “reasonable medical certainty” that medical experts often use at trial is not appropriate for SBS cases and can be extremely misleading. When medical experts cannot come to an agreement about SBS, a jury cannot understand the science well enough to make a reasonable decision.⁴³ Without corroborating **evidence** of child abuse, a defendant should not be convicted of any crime due to SBS based only on the existence of the triad of symptoms.

II. What Is **Shaken Baby Syndrome**?

The term SBS “evokes a powerful image of abuse.”⁴⁴ If doctors find certain symptoms in an infant that are characteristic of SBS, absent other *706 explanations for the injuries, then the last adult with the child is usually charged with intentional abuse.⁴⁵ After the case of British nanny Louise Woodward--who was convicted of shaking and killing a nine-month-old child in her care⁴⁶--captured media attention in 1997, child-abuse convictions increasingly interested the public.⁴⁷ There were roughly 104 reported appellate cases dealing with SBS in the United States in 2006.⁴⁸

Today, SBS is a popular phrase in the media, hospitals, and courtrooms.⁴⁹ People commonly associate it with a frustrated reaction to a crying baby.⁵⁰ Men are more likely to be accused: fathers or step-fathers account for the highest number of suspected abusers, and boyfriends of the baby's mother are the next highest.⁵¹ A rising number of babysitters are also being charged.⁵² There are approximately 1,500 reported cases of SBS nationwide every year.⁵³ In fact, according to the National Center on **Shaken Baby Syndrome**, the problem may be underreported because there is usually no “external **evidence** of trauma.”⁵⁴ The Center also estimated that 25 to 30 percent of shaken infants die from SBS-related injuries despite medical treatment.⁵⁵ Recent studies have indicated, however, that there may be other causes of the triad.⁵⁶

*707 A. Symptoms of **Shaken Baby Syndrome: The Triad**

In 1972, John Caffey first described SBS as “whiplash-shaking” of infants which he claimed resulted in bleeding in the brain and eyes without signs of impact to the head.⁵⁷ He later called this “whiplash shaken infant syndrome.”⁵⁸ Today,

SHAKEN BABY SYNDROME: MEDICAL UNCERTAINTY..., 2007 Wis. L. Rev. 701

SBS is characterized by a triad of symptoms: subdural or subarachnoid hematomas, retinal hemorrhaging, and brain swelling.⁵⁹ A subdural hematoma is a collection of blood between the outermost membrane of the brain (the dura mater) and the middle layer (the arachnoid).⁶⁰ A subarachnoid hematoma refers to the bleeding between the arachnoid layer and the innermost membrane (the pia mater).⁶¹ Retinal hemorrhaging is bleeding in the retina of the eyes from a ruptured blood vessel.⁶² Any one of these symptoms in isolation could come from another cause, but some doctors believe that a combination of the three only results from SBS.⁶³ Sometimes these injuries are accompanied by skull fractures, or broken or bruised ribs or arms.⁶⁴

In reality, SBS is more of a legal term than a medical term.⁶⁵ According to the SBS theory, the offender usually holds the baby by the torso or shoulders (which accounts for any fractures or bruises) and violently shakes the child back and forth.⁶⁶ This movement of the baby's head causes the veins connecting the brain to the skull to tear, leading to loss of oxygen to the brain and significant brain swelling⁶⁷ --which may *708 ultimately kill the baby.⁶⁸ A proper diagnosis requires an examination by many medical specialists, including a pathologist, a pediatrician, a neurologist, and an ophthalmologist.⁶⁹

B. Summary of Significant SBS Convictions

The earliest SBS convictions happened about twenty-five years ago,⁷⁰ and the number grows each year.⁷¹ The Woodward case was one of the most famous SBS cases, and the subject itself has been of growing interest to the media and society.⁷²

1. earliest cases

An early appellate case was *People v. Kailey*, in which the Supreme Court of Colorado affirmed Randy Steven Kailey's felony child-abuse conviction.⁷³ Kailey testified that, after picking up his four-month-old daughter from a babysitter and putting her to bed,⁷⁴ he woke up to find that she had vomited and was having trouble breathing.⁷⁵ Doctors found subdural hemorrhages, retinal hemorrhages, brain swelling, and bruises on her forehead and abdomen.⁷⁶ Four days later, surgery revealed that she had both an acute and a chronic subdural hematoma.⁷⁷ Two months earlier, the baby had been admitted to the hospital with a subdural hematoma, which the parents claimed came from her rolling off of the seat of the car or the couch.⁷⁸ At the trial, several doctors testified that the baby's injuries came from "either a blow to the head or whiplash shaken infant syndrome."⁷⁹ The defense unsuccessfully argued that the new injury was the result of a re-bleed from the earlier injury.⁸⁰

In another early appellate case, Janet Ostlund was convicted of second-degree murder for shaking her adopted daughter, Maria, who had *709 a history of health problems at the time of adoption.⁸¹ Ostlund claimed that, while alone with Maria, she turned her back, heard a thump, and then saw Maria on the floor near the couch.⁸² The baby died from brain swelling, and there was subdural bleeding.⁸³ The state's theory, based on circumstantial **evidence**, was that "a violent shaking" caused Maria's injuries.⁸⁴ Each side presented six expert witnesses, who provided the primary **evidence** in the case.⁸⁵ Despite the minimal and circumstantial **evidence** against Ostlund,⁸⁶ she was convicted of second-degree murder and sentenced to 105 months in prison.⁸⁷

2. commonwealth v. woodward

One of the most famous SBS cases was Commonwealth v. Woodward, in which Louise Woodward, a British nanny, was charged with murdering an eight-month-old boy.⁸⁸ A jury convicted Woodward of murder in the second degree, but the judge reduced the verdict to involuntary manslaughter and vacated her life sentence.⁸⁹ The judge concluded that Woodward did not act with malice-- an element of second-degree murder.⁹⁰ Additionally, the judge acknowledged the possibility of another cause for the SBS-type symptoms: the child had a “pre-existing skull fracture and blood clot”⁹¹ and Woodward's rough handling of the child caused the blood clot to “re-bleed.”⁹² Therefore, her actions were only “fatal because of [the child's] condition at the time.”⁹³ *710 Eight years later, the case remained in the media spotlight, as lawmakers, doctors, and parents passed legislation to prevent SBS.⁹⁴

C. The Changing Science

Today, there is no consensus among medical professionals as to whether the symptoms that have traditionally been attributed to SBS are necessarily indicative of intentional shaking.⁹⁵ New studies cast doubt on the conclusion that subdural hematomas and retinal hemorrhages in babies are definitive signs of SBS.⁹⁶ Many doctors have rejected these traditional notions and are looking to other causes--including falls, earlier trauma, and preexisting medical conditions.⁹⁷ In addition, studies have suggested that, if a child sustains an injury, there can be a “lucid interval” between the injury and the time of death,⁹⁸ making it difficult for doctors to establish when the injury occurred. Therefore, without corroborating **evidence**, it is harder to determine who, if anyone, might have inflicted the injury.

1. other causes of sbs symptoms

Studies have shown that findings of subdural hematomas and retinal hemorrhages are not always diagnostic of SBS.⁹⁹ A child may have preexisting disorders that can cause subdural hematomas,¹⁰⁰ including certain infections, clotting disorders, inherited disorders, coagulopathy (a *711 disease that affects the coagulation of blood), or re-bleeds of prior chronic hematomas.¹⁰¹ In addition, some studies have shown that short falls may cause acute subdural hematomas.¹⁰² Retinal hemorrhaging is another symptom that doctors often associate with SBS,¹⁰³ but it has been found in cases of accidental injury.¹⁰⁴ Subdural hematomas and retinal hemorrhages are considered primary injuries, which can then cause brain swelling.¹⁰⁵ Brain swelling “is not indicative of any specific telltale act, origin or cause” and a subdural hematoma can decrease oxygen to the brain which causes it to swell.¹⁰⁶ While doctors who believe in the traditional SBS theory agree that these symptoms alone may have other causes, they believe that the simultaneous existence of all three--without **evidence** of impact--is diagnostic of shaking.¹⁰⁷

Still, the possibility that other events may cause the SBS triad highlights the danger of relying solely on these symptoms to allege abuse.¹⁰⁸ This “raise[s] the possibility that virtually anyone could face years behind bars if, while in his or her care, a small child were to experience a devastating accident or the onset of an undetected illness whose symptoms resemble **shaken baby** syndrome's.”¹⁰⁹ Thus, prosecutors should not charge crimes based on SBS without corroborating **evidence**.

SHAKEN BABY SYNDROME: MEDICAL UNCERTAINTY..., 2007 Wis. L. Rev. 701

In addition, biomechanical research is casting doubt on traditional SBS theories. One recent report studied the biomechanics of shaking an infant.¹¹⁰ The researchers concluded that the infant neck could not withstand the forces commonly associated with SBS without whiplash injury;¹¹¹ furthermore, they found that shaking an infant with a force below the level traditionally associated with SBS would cause severe cervical-spinal-cord or brain-stem injury.¹¹² The study suggested that *712 because most SBS cases do not involve spine injury, a finding of SBS-like symptoms without a corresponding cervical-spine or brain-stem injury (which is common in many SBS cases) would imply the possibility of other causes for the injury.¹¹³ Other studies have also suggested that short falls can cause the triad by measuring the acceleration of a head falling onto various surfaces from short distances.¹¹⁴

Courts are becoming aware of the new scientific research surrounding SBS. In *Commonwealth v. Davis*, a Kentucky circuit court acknowledged that the traditional SBS theory is under debate and is not a certain science.¹¹⁵ The court concluded that the theory is accepted in the clinical medical community, but not necessarily in the general scientific community.¹¹⁶ In other words, the physicians who treat the babies “routinely diagnose SBS” when they observe the triad, but “this diagnosis is based on inconclusive research conducted in the scientific research community.”¹¹⁷ This is problematic because it amounts to the physician making a legal conclusion and not a medical opinion supported by science.¹¹⁸

2. the existence of lucid intervals

One can only identify the alleged perpetrator after determining when the injuries occurred. Magnetic resonance imaging is the most accurate method for determining the timing of injuries, but it can only provide a rough range of hours, at best.¹¹⁹ Furthermore, there is considerable controversy in the medical community as to whether a child can appear normal after sustaining subdural hematomas, retinal hemorrhaging, or massive brain swelling. Many doctors believe that a baby would not appear lucid after being shaken and would deteriorate rapidly after the injury.¹²⁰ Recent studies, however, have concluded that infants can have a “lucid interval” after suffering an injury that leads to *713 rapid deterioration.¹²¹ A 2005 study of children under four years of age suggested that, although it is rare, an infant or toddler can sustain a fatal head injury yet appear lucid to hospital staff before death.¹²² There was an overrepresentation of young children in this category: six of the children in the study¹²³ were lucid at admission, and five of those were less than two years old.¹²⁴ Four of the six lucid children “sustained a subdural hematoma as part of their head injury.”¹²⁵

One pathologist noted a case study in which a thirteen-month-old was “irritable, sleepy, and vomiting.”¹²⁶ The infant had difficulty breathing the next morning and doctors pronounced her brain dead that night.¹²⁷ Her autopsy showed subdural and retinal hemorrhages and brain swelling.¹²⁸ The symptoms from the severe brain injury were latent for several hours, and medical professionals did not notice anything particularly abnormal.¹²⁹

In 2000, an Illinois appellate court overturned an SBS conviction based on the timing of the infant's injuries.¹³⁰ In that case, while babysitting a sickly five-month-old, Donna Gist--a caretaker hired so that the parents could sleep--went to the bathroom and returned to find that the infant was not breathing.¹³¹ The autopsy showed that he had the classic triad of symptoms and bruising, but the medical examiner testified that she could only determine that his injuries occurred within twenty-four hours of his death.¹³² An expert witness for the prosecution testified that the injuries were inflicted

very close to the time of death because there was little brain swelling;¹³³ however, the defendant's doctor testified that she believed the infant sustained brain injury well before the defendant arrived that night.¹³⁴

*714 Gist was convicted of first-degree murder and initially sentenced to life in prison, but the trial court reduced the sentence to fifty years.¹³⁵ The appellate court reversed her conviction because more than one person had the opportunity to cause the injuries that killed the child.¹³⁶

III. The Role of Medical Experts in SBS Trials

“Beyond a reasonable doubt” and “reasonable medical certainty” are common phrases in criminal trials. The former refers to legal certainty in a conviction,¹³⁷ while the latter involves certainty in a medical diagnosis.¹³⁸ This difference has serious implications for SBS cases. Medical experts play an important role in these cases, but the testimony of these doctors often goes beyond the scope of medical experts in criminal trials.

A. Legal Standard Beyond a Reasonable Doubt

“Beyond a reasonable doubt” is the constitutional burden of persuasion by which the prosecution must prove “all the essential elements of guilt.”¹³⁹ While it is understood to be the standard in criminal law, it is unclear exactly what is meant by “reasonable doubt.”¹⁴⁰ This standard exists to protect innocent people from being convicted of crimes.¹⁴¹ The Supreme Court of the United States held that

[t]he accused during a criminal prosecution has at stake interests of immense importance, both because of the possibility that he may lose his liberty upon conviction and because of the certainty that he would be stigmatized by the conviction. Accordingly, a society that values the good name and freedom of every individual should not condemn a man for *715 commission of a crime when there is reasonable doubt about his guilt.¹⁴²

Massachusetts Supreme Judicial Court Chief Justice Lemuel Shaw once said in a jury instruction that the reasonable-doubt standard is far greater than the standard of “more likely than not,” and must convincingly establish the truth of the fact to a moral certainty.¹⁴³ Reasonable doubt was “the converse of the sum of both reasonable certainty and moral certainty,” and it was the doubt that a reasonable person would hold.¹⁴⁴ The juror did not need to specifically articulate the nature of the doubt to establish its reasonableness.¹⁴⁵ Courts, however, gradually began adopting jury instructions that required the juror to provide an articulated reason for the doubt.¹⁴⁶ In addition, courts slowly phased out the “moral certainty” language.¹⁴⁷ The ambiguity regarding moral certainty and the requirement of articulation dangerously shifted the standard for the jury by altering the presumption of innocence.¹⁴⁸

B. Reasonable Medical Certainty

While the jury must find guilt beyond a reasonable doubt, expert medical witnesses must base their opinions on a “reasonable medical certainty.”¹⁴⁹ Thus, during a physician's testimony, attorneys often ask if the physician can identify

SHAKEN BABY SYNDROME: MEDICAL UNCERTAINTY..., 2007 Wis. L. Rev. 701

the cause of an injury or death to a “reasonable medical certainty.”¹⁵⁰ Although this standard mostly applies to civil cases, doctors in criminal cases must also meet it.¹⁵¹ Unfortunately, the *716 meaning of this commonly used phrase is difficult to articulate:¹⁵² the “reasonable medical certainty” standard is probably not as high as the “beyond a reasonable doubt” standard,¹⁵³ which could be confusing in SBS cases.

Many lawyers assume that “reasonable medical certainty” is a medical term, but physicians only use it in litigation, and not in everyday practice.¹⁵⁴ Furthermore, physicians do not have one particular definition for the term¹⁵⁵: some consider it to be near the civil burden of proof of “more probable than not,”¹⁵⁶ while others consider it to be a “near absolute certainty,” closer to the higher criminal standard.¹⁵⁷

This confusion has led many physicians to apply their own understanding of the phrase to criminal cases. For example, in *Burke v. Town of Walpole*, the prosecution in a murder trial used a forensic odontologist who analyzed bite marks.¹⁵⁸ He used the phrase “reasonable degree of scientific certainty,” which he interpreted to be a “high degree of probability.”¹⁵⁹ He used those two terms interchangeably according to the “Bite-Mark Terminology Guidelines,” but clarified that he meant that there was “no reasonable or practical possibility that someone” other than the defendant made the bite mark.¹⁶⁰ In other words, the bite-mark guidelines equated “reasonable medical certainty” with a high standard similar to “beyond a reasonable doubt.”¹⁶¹ Despite the higher standard that the guidelines intended, the First Circuit Court of Appeals used a lower “probable cause” standard in that case.¹⁶²

*717 Because the term is amorphous, medical experts have wide discretion in testifying about issues of probability.¹⁶³ To make matters worse, courts and attorneys have not come to a consensus on what the commonly used phrase means.¹⁶⁴ Therefore, not only are testifying doctors suggesting their own interpretations,¹⁶⁵ but the attorneys examining them and the judges and jurors interpreting the **evidence** may also have different understandings of the meaning of “reasonable medical certainty.”

Commentators have suggested that the term seeks to permit medical experts to give their opinions without an absolute certainty, so as not to impose on the fact-finder's role.¹⁶⁶ The phrase has expanded beyond its original intention, and most states have incorporated it into both civil and criminal statutes.¹⁶⁷ Experts have even used the term “medical certainty” in DNA cases,¹⁶⁸ which is troubling. DNA testing does not provide a certain match to an individual, but instead gives “the statistical probability that a person picked randomly from the population would have a DNA profile identical to the DNA profile generated from the forensic sample.”¹⁶⁹ In *Howard v. State*, the Supreme Court of Mississippi noted that an expert testifying about DNA **evidence** must make this distinction clear.¹⁷⁰ The court questioned the validity of a bite-mark expert's claim that the bite marks found on the victim matched the defendant to a “reasonable medical certainty,” when even DNA experts could not make such a claim.¹⁷¹

The lack of a stable definition has grave consequences because doctors commonly use the phrase when testifying in SBS cases.¹⁷² If they *718 are using it as the equivalent of the “more likely than not” standard (or the First Circuit's even lower probable-cause standard), however, that should not be enough to sustain a conviction.¹⁷³ Even if these doctors equate the term with a higher standard, the questionable validity of the triad makes medical certainty a difficult standard to reach.¹⁷⁴ In an SBS case based only on medical testimony, such testimony should make it clear to the jury that there cannot be absolute certainty.

C. Role of Doctors in SBS Cases

The prosecution's doctors usually testify that the child had the traditional triad of symptoms, and explicitly state that only SBS could have caused the injuries.¹⁷⁵ They also estimate a time range in which the injury occurred.¹⁷⁶ If a defendant attributes the baby's injuries to an accidental fall,¹⁷⁷ the prosecution's witness usually refutes that argument by stating that the triad of injuries could only come from a violent shaking.¹⁷⁸ Recent studies, however, suggest that short falls can cause SBS-like symptoms.¹⁷⁹ It is not appropriate for medical experts to guess what specifically occurred when the child was injured.

Sometimes, in estimating when the child sustained its injuries, doctors look to information other than medical findings.¹⁸⁰ For example, they may incorporate witness statements about the child's condition before the injury,¹⁸¹ and may consider whether the alleged perpetrator *719 has a history of being abusive or experienced abuse as a child.¹⁸² In addition, in reporting their findings to the prosecutor, these experts may already know the prosecution's main suspect and may review police reports before making a final conclusion.¹⁸³ The potential bias that could result is similar to a psychologist's "examiner bias,"¹⁸⁴ which can occur in SBS cases when doctors examine information outside of the medical records and form a hypothesis about the perpetrator of the assumed abuse before reporting their findings.¹⁸⁵

Medical experts should be extremely careful in diagnosing SBS in criminal cases,¹⁸⁶ and should limit their examination to medical findings when making conclusions about the cause of death. This line is difficult to draw because doctors may need to extract information from the caregiver or the police to treat the child effectively. Such information, however, should not be used in forming expert opinions about the cause of the injuries.

It is particularly hard to obtain quality **evidence** in SBS cases.¹⁸⁷ According to critics of the traditional SBS theory, doctors may want to protect children, but they should not provide opinions that are not medically sound.¹⁸⁸ Experts need to be aware of other conditions that may mimic child abuse,¹⁸⁹ and they should rule out every other possibility before making any accusations.¹⁹⁰ The role of an expert is to assist the trier of fact in understanding the **evidence**;¹⁹¹ in SBS cases, their role should be to state their medical findings and allow the jury to *720 find the facts¹⁹²--particularly the identification of the alleged perpetrator and conclusions about whether a crime actually occurred.

By using the triad of symptoms to diagnose SBS, the "physician is diagnosing the legal conclusion that someone has battered [a] child without manifest signs of bruising, broken bones, or other **evidence**."¹⁹³ Furthermore, the conclusions are circular: SBS is diagnosed based on the triad of symptoms because prior defendants have been convicted under an SBS theory when a child presents with the triad.¹⁹⁴

IV. Looking to Other Countries for Solutions

Courts in other countries are also dealing with SBS cases and the emergence of conflicting science. For example, the courts in England have reexamined SBS convictions that were supported only by medical.¹⁹⁵ In *R v. Cannings*, the Court of Appeals Criminal Division--one of the highest courts in England-- overturned Angela Cannings's conviction of killing two of her children by smothering.¹⁹⁶ After reading expert reports, the court concluded "that a great deal about

SHAKEN BABY SYNDROME: MEDICAL UNCERTAINTY..., 2007 Wis. L. Rev. 701

death in infancy, and its causes, remains as yet unknown and undiscovered.”¹⁹⁷ The Court held that, when there is conflicting medical testimony presented at trial, the jury cannot proceed without additional **evidence**.¹⁹⁸ Following this judgment, the Attorney General set up a group to review convictions in alleged SBS cases.¹⁹⁹ As a result, the committee considered 297 cases.²⁰⁰

The Court then heard four appeals together,²⁰¹ and found that “[t]he common thread running through each of these four appeals is a submission that since these convictions medical research has developed to the extent that there is now ‘fresh **evidence**’ which throws doubt on *721 the safety of each conviction.”²⁰² In *R v. Harris*, the Court quashed Loraine Harris's manslaughter conviction.²⁰³ Although the experts for the state asserted that he died from being shaken,²⁰⁴ the Court concluded that “the mere presence of the triad on its own cannot automatically or necessarily lead to a diagnosis of [SBS].”²⁰⁵ In *R v. Faulder*, Michael Faulder was convicted of causing grievous bodily harm to a seven-week-old boy.²⁰⁶ He claimed that he had accidentally dropped the baby.²⁰⁷ There were no retinal hemorrhages, but there were subdural hemorrhages and brain swelling.²⁰⁸ The infant made a full recovery,²⁰⁹ and the Crown ended up changing its case from arguing that Faulder shook the infant to arguing that he delivered multiple blows to its head.²¹⁰ The Court quashed the conviction in part because defense experts proposed alternative explanations for the injury.²¹¹

These appeals tested the reliability of expert **evidence** used in SBS cases, and “could lead to a rethink in the way such cases are treated.”²¹² The Court examined the triad, the degree of force necessary to cause SBS-like symptoms, and biomechanics involved with shaking infants,²¹³ and reiterated that changes in science “should not be kept from the Court.”²¹⁴ It went on to stress the obligations of expert witnesses, including that the expert “should never assume the role of an advocate.”²¹⁵ In the end, the Court noted that it depended on medical witnesses to provide information on the issues involved in these types of cases.²¹⁶ The *London Times* predicted that there would be demands in the future that courts analyze SBS cases individually on the **evidence**, and not on the questionable SBS triad.²¹⁷

The Court recognized that the triad should not always lead to a conclusion of SBS, and that courts should examine the facts on a case- *722 by-case basis.²¹⁸ Graham Zellick, the chair of England's Criminal Cases Review Commission, has argued that the medical **evidence** is too complicated for juries in SBS cases, leading to the convictions of innocent people.²¹⁹ He has suggested that the judge hear the medical **evidence** away from the jury and then direct the jury on what to make of the **evidence**.²²⁰ This is similar to judicial direction of juries on the application of law.²²¹ SBS cases are unique in this regard because of the complexity of the science and the medical debate.²²²

The Supreme Court of Western Australia has also recognized the controversy surrounding the SBS triad and convictions without corroborating **evidence**.²²³ In one case, a baby exhibited the constellation of symptoms associated with SBS in addition to other injuries.²²⁴ The prosecution argued that the defendant had violently shaken the child to death.²²⁵ Professor John Hilton, an expert of pathology, testified for the defense that the proposition that only SBS can cause such injuries “is highly suspect in any individual case unless there is a reliable witness.”²²⁶ The court interpreted this to mean that, without an eyewitness, **evidence** of external injuries such as bruising or a confession, a conclusion of SBS is “highly suspect.”²²⁷ The trial judge held that he could not conclude beyond a reasonable doubt that the child had been shaken to death,²²⁸ and the appellate court agreed.²²⁹ There have been several other recent Australian cases in which suspected child abusers were acquitted.²³⁰

SHAKEN BABY SYNDROME: MEDICAL UNCERTAINTY..., 2007 Wis. L. Rev. 701

Canadian courts are also permitting changes in science to influence their decisions. In one case, a child died of brain injury in Ontario.²³¹ Although there was no immediate autopsy, Charles Smith, a pediatric pathologist, exhumed the body three weeks later, conducted an autopsy, and concluded that a twelve-year-old babysitter had killed the baby by shaking it.²³² Smith, however, did not follow “basic procedures for arriving at his conclusion,”²³³ and also ignored a deep bruise on the forehead, which corroborated the babysitter's claim of an accidental fall.²³⁴ At trial, Smith stated that the baby had to have been shaken to death because it could not have died from a short fall.²³⁵ The twelve-year-old was eventually acquitted,²³⁶ and the judge advised that Smith stay abreast of the current research on SBS, explaining that a doctor should always consider possibilities other than shaking.²³⁷

An independent team of experts is thoroughly examining forty homicides and suspicious deaths that Smith investigated at the Hospital for Sick Children since 1991.²³⁸ Until recently, Smith was a powerful medical expert in Ontario, who bragged about getting more convictions than other experts “against child killers.”²³⁹ He was a persuasive witness and was Ontario's “top forensic expert on suspicious child deaths” for over a decade.²⁴⁰ Now, judges and medical authorities have criticized him for jumping to conclusions and for tardy reporting.²⁴¹ There is significant doubt about his conclusions in more than one thousand autopsies.²⁴²

Finally, researchers in Hong Kong and Japan are also taking a careful look at SBS cases.²⁴³ In Hong Kong, if there is suspicion of SBS--regionally referred to as nonaccidental injury--then pediatricians, social workers, police, schoolteachers, and other medical specialists hold a case conference to discuss the findings and decide whether the baby was shaken.²⁴⁴ They also create a plan for its future welfare.²⁴⁵ In one such instance, a child showed acute subdural hematomas and retinal hemorrhages, which appeared to be caused by SBS.²⁴⁶ The family provided “good social support and the mother was mature and emotionally stable,” and did not seem suspicious.²⁴⁷ The members of the conference debated about whether falling onto a mattress could have produced such symptoms, and concluded that the cause was not SBS.²⁴⁸ If the mother had been a drug addict or a single parent, the conclusion might have been different:²⁴⁹ “[y]oung parents, unstable family situations, low socioeconomic status and disability of the child are well known risk factors for [SBS].”²⁵⁰

A Japanese study concluded that retinal hemorrhages and a subdural hematoma with no external signs of injury were usually attributed to accidental or trivial head injury, while subdural hemorrhages associated with signs of external trauma to the face or head were commonly found in child-abuse cases.²⁵¹ Additionally, subdural hematomas and retinal hemorrhages may develop from minor injuries or “casual shaking.”²⁵² According to some commentators, Hong Kong's method of holding conferences²⁵³ in cases of unexplained subdural hematomas--even without a history of abuse--may “run the risk of damaging the lives of innocent families, in [an] attempt to prevent further injury to this or other children.”²⁵⁴

*725 V. Conclusion

There is no doubt that child abuse exists and that the perpetrators must be charged with these crimes. Violent shaking may cause serious injury or death in infants, but the legal system should not convict people based exclusively on disputed

SHAKEN BABY SYNDROME: MEDICAL UNCERTAINTY..., 2007 Wis. L. Rev. 701

medical **evidence**. If there is no corroborating **evidence**--such as an eyewitness testimony, a confession,²⁵⁵ a history of abuse, or external injuries--then such disputed medical **evidence** should not overcome a reasonable doubt.

There is growing debate in the medical community as to whether the triad may have other causes. Jennian Geddes has suggested that the triad should not be called SBS at all, but “infantile encephalopathy with subdural and retinal bleeding.”²⁵⁶ This alternative name accurately describes the medical condition, and does not imply that someone committed a violent act.²⁵⁷ SBS can invoke an emotional response because of the tragedy of an infant's death.²⁵⁸ Once jurors hear the term “**shaken baby syndrome**” in the courtroom, they may be quick to jump to conclusions about the defendant.

Like Smith, experts for the prosecution in U.S. SBS cases likely consider themselves to be advocates for child victims. SBS cases illustrate the emotional controversy, because “child protection [is] a field in which emotion often threatens to overwhelm scientific objectivity.”²⁵⁹ One defense lawyer in Australia tried to retain medical experts for a trial and an ophthalmologist refused to help her, stating “I'm on the baby's side.”²⁶⁰ Some doctors testify for the prosecution in many SBS cases, possibly pursuing their own agendas instead of basing their opinions solely on objective medical **evidence**.²⁶¹ Any such conclusions about how the baby received the injuries or who caused them could greatly prejudice the jury.

Experts should also be prohibited from looking at nonmedical documents. **Rule 703 of the Federal Rules of Evidence** states that an expert may base an opinion on facts or data available at or before the *726 hearing.²⁶² The purpose of this rule is to broaden the basis for opinions,²⁶³ so keeping experts from considering police reports or stories from the caregiver might not be feasible.²⁶⁴ Doctors should not rely on such **evidence**, however, in making their conclusions.

Ultimately, doctors should not use the term “**shaken baby syndrome**” in courts at all, because it conjures up an image of abuse and violence and merely represents a legal conclusion that attempts to describe what happened to the baby. Experts cannot testify that the child died from SBS--or when it was shaken--with “reasonable medical certainty,” particularly when there is no consensus in the medical community. Instead, medical experts should only describe the baby's symptoms and medical history, and clarify that a conclusion cannot be made to a “reasonable medical certainty.”

Recognizing that the triad is not necessarily diagnostic of SBS, the Court of Appeal of England wisely reexamined SBS cases in which there was no corroboration of the medical **evidence**.²⁶⁵ Charges of SBS must have corroboration because the triad of symptoms cannot meet the high standard of “beyond a reasonable doubt.” Courts in some other countries are realizing the risk of **wrongful convictions** and taking important steps to remedy the problem by examining other causes of the symptoms and requiring corroborating **evidence**.

Prosecutors, therefore, should not charge anyone with a crime based on SBS without corroborating **evidence**. The existence of the triad of symptoms alone is insufficient to support a conviction, and should not meet the “probable cause” standard.²⁶⁶ The public concern with infant deaths, however, may pressure prosecutors to continue charging in these cases.²⁶⁷

In addition, trial judges should dismiss cases that do not have corroborating **evidence**. Again, this is probably not feasible because of the tragedy involved and the public's interest in holding people *727 accountable for infant deaths. In England, while Zellick's idea of having only the judge hear the medical **evidence**²⁶⁸ is appealing, a judge is not necessarily in a better position than the jury to understand such disputed and complicated medical **evidence**.

Finally, and perhaps most feasibly, given the current climate, state legislatures should set up taskforces to examine supposed SBS cases more thoroughly. Instead of replicating Hong Kong's case conferences, implementing an independent taskforce of doctors, scientists, judges, lawyers, and child-abuse specialists would allow examination of all arguments and establishment of standards for cases in which the baby only suffered from the triad of symptoms. This taskforce could look to other countries and studies from various medical fields--including pathology, neurology, ophthalmology, pediatrics, and biomechanics--in coming to a conclusion about how to handle such cases. Because medical professionals disagree on the merits of the SBS triad and whether such injuries can be accurately timed, courts must adapt to changes in the medical community.²⁶⁹ A taskforce should study **evidence**-based SBS research²⁷⁰ and legal challenges, and create a standard for "reasonable medical certainty" to help prevent **wrongful convictions** in already tragic cases of infant deaths.

Footnotes

a1 J.D., University of Wisconsin Law School. I dedicate this Comment to Audrey Edmunds, who is currently serving an eighteen-year sentence for allegedly shaking a baby to death in 1995. I worked on her case with the Wisconsin Innocence Project, and her story inspired this Comment. I would like to thank Shelley Fite, Katie Mason, Keith Findley, and John Pray for their guidance. I would also like to thank all of the doctors and lawyers who took the time to speak with me, especially John Plunkett, Donna Kuchler, William Perloff, and Shaku Teas. Finally, thanks to Rhoda, Peter, Lesley, Rachel, Adam, David, and Owen for putting up with me through the research and writing of this Comment.

1 R v. Cannings, [2004] EWCA (Crim) 1, ¶ 179, 2 Crim. App. 7 (H.L.) (appeal taken from Winchester).

2 Steve Sharp, Dittberner Jury Is Selected, Watertown Daily Times (Watertown, Wis.), June 21, 2005, at 1.

3 Watertown Man Accused of Shaking to Death Son Is Found Not Guilty, Duluth News Trib., July 11, 2005 [hereinafter Watertown Man Accused].

4 Id.

5 Sharp, *supra* note 2.

6 See When Babies Cry, 70 Tex. B.J. 409, 409 (2007); Alan Mayor Sokobin, **Shaken Baby Syndrome: A Comparative Study: Anglo-American Law and Jewish Law--Legal, Moral, and Ethical Issues**, 29 U. Tol. L. Rev. 513, 518 (1998)

7 Telephone Interview with Donna Kuchler, Todd Dittberner's Defense Attorney (Jan. 5, 2006).

8 Id.

9 Steve Sharp, Eye Expert Testifies for Dittberner, Watertown Daily Times (Watertown, Wis.), June 30, 2005, at 1.

10 Watertown Man Accused, *supra* note 3.

11 Sharp, *supra* note 9.

12 Id.

13 Telephone Interview with Donna Kuchler, *supra* note 7.

14 Id.

15 Sharp, *supra* note 2.

SHAKEN BABY SYNDROME: MEDICAL UNCERTAINTY..., 2007 Wis. L. Rev. 701

- 16 Id.
- 17 Id.
- 18 Sharp, *supra* note 9.
- 19 Id.
- 20 Id.
- 21 Id.
- 22 Id.
- 23 Id.
- 24 Id.
- 25 Telephone Interview with Donna Kuchler, *supra* note 2.
- 26 Id.
- 27 Id. Daycare workers also testified that Riley had trouble with his lungs, and that they were concerned because he often struggled to breathe. Sharp, *supra* note 9.
- 28 Telephone Interview with Donna Kuchler, *supra* note 7.
- 29 Id.
- 30 Watertown Man Accused, *supra* note 3.
- 31 While retaining experts may be prohibitively expensive in some cases, another problem for defendants is that some doctors might be afraid to testify. One defense expert, John Plunkett--who has testified in more than 100 SBS cases--was charged with false swearing based on his testimony that the traditional triad of symptoms is not necessarily diagnostic of SBS. Mark Hansen, *Battle of the Expert*, A.B.A. J., Dec. 2005, at 52, 54.
- 32 See, e.g., *State v. Ostlund*, 416 N.W.2d 755, 760 (Minn. Ct. App. 1987).
- 33 See Richard Guilliat, *When the Bough Breaks*, *Good Weekend*, Nov. 20, 2004, at 18, 19.
- 34 See Joseph D. Hatina, Note, *Shaken Baby Syndrome: Who Are the True Experts?*, 46 *Clev. St. L. Rev.* 557, 567 (1998).
- 35 See Guilliat, *supra* note 33.
- 36 See Charles A. Phipps, *Responding to Child Homicide: A Statutory Proposal*, 89 *J. Crim. L. & Criminology* 535, 537 (1999).
- 37 See Sharp, *supra* note 9; Telephone Interview with Donna Kuchler, *supra* note 7.
- 38 See *R v. Cannings*, [2004] EWCA (Crim) 1, ¶ 7, 2 *Crim. App.* 7 (H.L.) (appeal taken from Winchester) (“In the vast majority of cases of murder, there is no doubt that someone has caused or contributed to the death of the deceased in some way Unusually, but not uniquely, the primary and stark question in [this infant-death case] was whether either of these children was killed at all.”) Although *Cannings* was not an SBS case, it did involve infant deaths. See *id.* ¶ 9.
- 39 See Lee Scheier, *Shaken Baby Syndrome: A Search for Truth*, *Chi. Trib.*, June 12, 2005, *Magazine*, at 10, 12.
- 40 See *id.*

SHAKEN BABY SYNDROME: MEDICAL UNCERTAINTY..., 2007 Wis. L. Rev. 701

- 41 Matthew D. Ramsey, [A Nuts and Bolts Approach to Litigating the Shaken Baby or Shaken Impact Syndrome](#), 188 Mil. L. Rev. 1, 3 (2006).
- 42 See *infra* Part II.C.
- 43 See Clare Dyer, “Innocent Parents at Risk” in *Baby Death Cases Too Complex for Juries*, *Guardian* (London), Nov. 7, 2005, at 15.
- 44 J.F. Geddes & J. Plunkett, Editorial, The **Evidence** Base for **Shaken Baby** Syndrome, 328 *BMJ* 719, 719 (2004).
- 45 Genie Lyons, Comment & Note, **Shaken Baby** Syndrome: A Questionable Scientific Syndrome and a Dangerous Legal Concept, 2003 *Utah L. Rev.* 1109, 1110 (2003); see also Patrick D. Barnes, Ethical Issues in Imaging Nonaccidental Injury: Child Abuse, 13 *Topics in Magnetic Resonance Imaging* 85, 86 (2002).
- 46 *Commonwealth v. Woodward*, 694 N.E.2d 1277, 1281 (Mass. 1998).
- 47 See Phipps, *supra* note 36, at 536.
- 48 This figure is based on a March 5, 2007 LexisNexis search for “‘shaken baby syndrome’ OR ‘shaken infant syndrome’ OR ‘shaken impact syndrome’” performed on federal and state cases reported within the previous year.
- 49 See R. Uscinski, **Shaken Baby** Syndrome: Fundamental Questions, 16 *British J. Neurosurgery* 217, 217 (2002).
- 50 See Sokobin, *supra* note 6, at 518; see also Michael Levenson, *Lawmakers Target Shaken Baby Syndrome*, *Boston Globe*, Oct. 27, 2005, at B1 (“It’s not uncommon to feel extremely frustrated when you’re trying to care for a crying baby, so much so that probably everybody has had the thought about shaking a baby But the important thing is that you don’t carry out the act.”).
- 51 See Levenson, *supra* note 50.
- 52 See Sokobin, *supra* note 6, at 518.
- 53 Levenson, *supra* note 50.
- 54 National Center on **Shaken Baby** Syndrome, <http://www.dontshake.com/> (last visited July 17, 2007).
- 55 *Id.*
- 56 See, e.g., Barnes, *supra* note 45, at 86; Mark Donohoe, **Evidence**-Based Medicine and **Shaken Baby** Syndrome: Part I: Literature Review, 1966-1998, 24 *Am. J. Forensic Med. & Pathology* 239, 240-41 (2003); Geddes & Plunkett, *supra* note 44, at 719; Lyons, *supra* note 45, at 1120-25.
- 57 See John Caffey, *On the Theory and Practice of Shaking Infants*, 124 *Am. J. Diseases & Children* 161, 161, 169 (1972).
- 58 See John Caffey, *The Whiplash Shaken Infant Syndrome: Manual Shaking by the Extremities with Whiplash-Induced Intracranial and Intraocular Bleedings, Linked with Residual Permanent Brain Damage and Mental Retardation*, 54 *Pediatrics* 396, 396 (1974).
- 59 See, e.g., John Plunkett, **Shaken Baby** Syndrome and the Death of Matthew Eappen, 20 *Am. J. Forensic Med. & Pathology* 17, 17 (1999); Sam Lister, Q & A: **Shaken Baby** Syndrome, *Times Online*, Feb. 14, 2006, <http://www.timesonline.co.uk/tol/news/uk/article546383.ece>; Letter from William H. Perloff to Troy Cross, Assistant District Attorney, Columbia County, Wis., at 3 (Jan. 9, 2005) (on file with the Wisconsin Law Review).
- 60 Lyons, *supra* note 45, at 1110.
- 61 *Id.*

SHAKEN BABY SYNDROME: MEDICAL UNCERTAINTY..., 2007 Wis. L. Rev. 701

- 62 Hatina, *supra* note 34, at 559 n.15.
- 63 *Id.*
- 64 Phipps, *supra* note 36, at 545.
- 65 See James Le Fanu, *Wrongful Diagnosis of Child Abuse--A Master Theory*, 98 *J. Royal Soc'y Med.* 249, 251 (2005); Robert Tharp, *Shaken-Baby Theory Questioned*, *Dallas Morning News*, Mar. 10, 2006, at 2B (referring to SBS as a “law enforcement theory”).
- 66 Phipps, *supra* note 36, at 543.
- 67 *Id.* at 544.
- 68 See *id.*
- 69 Hatina, *supra* note 34, at 570 n.109.
- 70 See *infra* Part II.B.1.
- 71 See Lyons, *supra* note 45, at 1112-13.
- 72 See, e.g., Levenson, *supra* note 50; Kieran Nicholson, *Toddler Shaken, Doctor Testifies*, *Denver Post*, Sept. 8, 2000, at B2; Scheier, *supra* note 39.
- 73 662 P.2d 168 (Colo. 1983).
- 74 *Id.* at 169.
- 75 *Id.*
- 76 *Id.*
- 77 See *id.* at 170. An acute hematoma is sudden, while a chronic hematoma can develop over a period of time. See Lyons, *supra* note 45, at 1110-11.
- 78 Kailey, 662 P.2d at 170.
- 79 *Id.*
- 80 *Id.*
- 81 *State v. Ostlund*, 416 N.W.2d 755, 757 (Minn. Ct. App. 1987).
- 82 *Id.*
- 83 *Id.*
- 84 *Id.* at 758.
- 85 See *id.* at 758-61. There were no witnesses to the incident, *id.* at 757, although several relatives testified that Ostlund had lightly shaken the child on previous occasions. See *id.* at 763.
- 86 The Ostlund court held that “[a] conviction may be based on circumstantial evidence and will be upheld if the reasonable inferences from such evidence are consistent only with the defendant's guilt and inconsistent with any rational hypothesis except that of guilt.” *Id.* at 758 (citing *State v. Anderson*, 379 N.W.2d 70, 75 (Minn. 1986)).

SHAKEN BABY SYNDROME: MEDICAL UNCERTAINTY..., 2007 Wis. L. Rev. 701

- 87 [In re Welfare of M.D.O.](#), 450 N.W.2d 655, 656 (Minn. Ct. App. 1990) (summarizing the Ostlund decision).
- 88 694 N.E.2d 1277, 1281 (Mass. 1998); see also Phipps, *supra* note 36, at 536; Scheier, *supra* note 39, at 12.
- 89 Woodward, 694 N.E.2d at 1281.
- 90 See Phipps, *supra* note 36, at 537 (quoting [Commonwealth v. Woodward](#), 7 Mass. L. Rep. 449, 452, No. Crim. 97-0433, 1997 WL 69114 at *6 (Super. Ct. 1997)).
- 91 See Woodward, 694 N.E.2d at 1287. Interestingly, the Kailey court had earlier rejected this theory. See [People v. Kailey](#), 662 P.2d 168, 170 (Colo. 1983).
- 92 Woodward, 694 N.E.2d at 1287.
- 93 *Id.*
- 94 See [Mass. Gen. Laws ch. 111, § 24K](#) (2006); see Levenson, *supra* note 50. This law aims to teach parents how to calm a crying baby without resorting to shaking. See *id.*
- 95 See, e.g., Donohoe, *supra* note 56, at 241; Lyons, *supra* note 45, at 1111-12; Scheier, *supra* note 39, at 12.
- 96 See, e.g., Eva Lai Wah Fung et al., Unexplained Subdural Hematoma in Young Children: Is It Always Child Abuse?, 44 *Pediatrics Int'l* 37, 37 (2002); Scheier, *supra* note 39, at 12; Lyons, *supra* note 45, at 1111.
- 97 Scheier, *supra* note 39, at 12.
- 98 See, e.g., Kristy B. Arbogast et al., Initial Neurologic Presentation in Young Children Sustaining Inflicted and Unintentional Fatal Head Injuries, 116 *Pediatrics* 180, 181 (2005); Scott Denton & Darinka Mileusnic, Delayed Sudden Death in an Infant Following an Accidental Fall: A Report with Review in the Literature, 24 *Am. J. Forensic Med. & Pathology* 371, 371 (2003); Mary G. Gilliland, Interval Duration Between Injury and Severe Symptoms in Nonaccidental Head Trauma in Infants and Young Children, 43 *J. Forensic Sci.* 723, 723 (1998).
- 99 See, e.g., Donohoe, *supra* note 56, at 241; Lyons, *supra* note 45, at 1111-12; Scheier, *supra* note 41, at 12
- 100 See, e.g., Mark Donohoe, [Shaken Baby Syndrome \(SBS\) and Non-Accidental Injuries \(NAI\)](#), Aug. 20, 2001, § 1.2.1, available at <http://www.whale.to/v/sbs.html> (giving a “framework for assessing assumptions and **evidence**” regarding SBS).
- 101 See *id.*; see also Lyons, *supra* note 45, at 1111.
- 102 See Barnes, *supra* note 45, at 86. See generally John Plunkett, Fatal Pediatric Head Injuries Caused by Short-Distance Falls, 22 *Am. J. Forensic Med. & Pathology* 1 (2001).
- 103 See Hatina, *supra* note 34, at 566.
- 104 P.E. Lantz et al., Perimacular Retinal Folds from Childhood Head Trauma, 328 *BMJ* 754, 756 (2004).
- 105 Ramsey, *supra* note 41, at 11.
- 106 *Id.* at 12.
- 107 See, e.g., Letter from William H. Perloff to Troy Cross, *supra* note 59, at 3.
- 108 J.F. Geddes et al., Dural Haemorrhage in Non-Traumatic Infant Deaths: Does It Explain the Bleeding in ‘[Shaken Baby Syndrome](#)’?, 29 *Neuropathology & Applied Neuropathology* 14, 20 (2003).
- 109 Scheier, *supra* note 39, at 12.

SHAKEN BABY SYNDROME: MEDICAL UNCERTAINTY..., 2007 Wis. L. Rev. 701

- 110 Faris A. Bandak, **Shaken Baby** Syndrome: A Biomechanics Analysis of Injury Mechanisms, 151 *Forensic Sci. Int'l* 71, 72 (2005).
- 111 *Id.* at 78.
- 112 *Id.*
- 113 *Id.*
- 114 See Scheier, *supra* note 39, at 26.
- 115 No. 04-CR-205, at 21 (Greenup County Cir. Ct. Apr. 17, 2006).
- 116 See *id.* at 23.
- 117 *Id.* at 22.
- 118 See *id.* at 23.
- 119 Barnes, *supra* note 45, at 89.
- 120 Marcus B. Nashelsky & Jay D. Dix, The Time Interval Between Lethal Infant Shaking and Onset of Symptoms: A Review of the **Shaken Baby** Syndrome Literature, 16 *Am. J. Forensic Med. & Pathology* 154, 154 (1995) (reviewing medical literature that expresses this view and mentioning the lack of study into the length of the interval).
- 121 See, e.g., Arbogast et al., *supra* note 98, at 181; Denton & Mileusnic, *supra* note 98, at 374.
- 122 Arbogast et al., *supra* note 98, at 184.
- 123 The study included 314 children: 37 percent sustained inflicted injuries, 13 percent fell, and 49 percent were in motor-vehicle crashes. *Id.* at 181. Therefore, the study did not exclusively deal with SBS cases.
- 124 *Id.*
- 125 *Id.*
- 126 Robert W. Huntington, III, Letter to the Editor, Symptoms Following Head Injury, 23 *Am. J. Forensic Med. & Pathology* 105 (2002).
- 127 *Id.*
- 128 *Id.*
- 129 *Id.*
- 130 *State v. Gist*, No. 2-98-1103 (Ill. App. Ct. May 2, 2000).
- 131 *Id.* at 7.
- 132 See *id.* at 9-11.
- 133 *Id.* at 10-11.
- 134 *Id.* at 12-13.
- 135 *Id.* at 14.

SHAKEN BABY SYNDROME: MEDICAL UNCERTAINTY..., 2007 Wis. L. Rev. 701

- 136 Id. at 15. In addition, the court noted that, even if it believed the doctor who placed the infliction of the injuries closer to the death, the state did not establish that the parents could not have committed the crime; they were only fourteen steps away. Id. at 17.
- 137 *In re Winship*, 397 U.S. 358, 361 (1970).
- 138 See Jeff L. Lewin, *The Genesis and Evolution of Legal Uncertainty about "Reasonable Medical Certainty,"* 57 Md. L. Rev. 380, 382 (1998)
- 139 *In re Winship*, 397 U.S. at 361-62.
- 140 See Steve Sheppard, *The Metamorphoses of Reasonable Doubt: How Changes in the Burden of Proof Have Weakened the Presumption of Innocence*, 78 Notre Dame L. Rev. 1165 (2003).
- 141 See *id.* at 1166.
- 142 *Winship*, 397 U.S. at 363-64.
- 143 See Sheppard, *supra* note 140, at 1200 (quoting *Commonwealth v. Webster*, 59 Mass. (5 Cush.) 295, 320 (1850)).
- 144 Id. at 1201.
- 145 Id.
- 146 Id. at 1210-11 (citing *State v. Jefferson*, 10 So. 199, 200 (La. 1891)).
- 147 See *Victor v. Nebraska*, 511 U.S. 1, 16-17 (1994) ("At the same time, however, we do not condone the use of the phrase [moral certainty]. As modern dictionary definitions of moral certainty attest, the common meaning of the phrase has changed since it was used in the Webster instruction, and it may continue to do so to the point that it conflicts with the Winship standard. Indeed, the definitions of reasonable doubt most widely used in the federal courts do not contain any reference to moral certainty."); see also Sheppard, *supra* note 140, at 1222-23.
- 148 See Sheppard, *supra* note 140, at 1239-40 ("The defendant in the face of the modern reasonable-doubt instruction must convince the jury of his innocence.").
- 149 See Lewin, *supra* note 138, at 382; see also *Addington v. Texas*, 441 U.S. 418, 430 (1979) (deeming this the appropriate standard). Courts and witnesses use the term "reasonable medical certainty" interchangeably with "reasonable degree of medical certainty." Id.
- 150 See Lewin, *supra* note 138, at 418-20.
- 151 See Barnes, *supra* note 45, at 90.
- 152 See Lewin, *supra* note 138, at 398; Glenn E. Bradford, *Dissecting Missouri's Requirement of "Reasonable Medical Certainty,"* 57 J. Mo. Bar 136, 142 (2001). Black's Law Dictionary defines "reasonable medical certainty" as "[i]n proving the cause of an injury, a standard requiring a showing that the injury was more likely than not caused by a particular stimulus, based on the general consensus of recognized medical thought." Black's Law Dictionary 1294 (8th ed. 2004).
- 153 See Bradford, *supra* note 152, at 141.
- 154 Lewin, *supra* note 138, at 402; Bradford, *supra* note 152, at 137-38.
- 155 Lewin, *supra* note 138, at 402; Bradford, *supra* note 152, at 141 ("It apparently means whatever the testifying physician thinks it means.").

SHAKEN BABY SYNDROME: MEDICAL UNCERTAINTY..., 2007 Wis. L. Rev. 701

- 156 Barnes, *supra* note 45, at 90 (“The testimony offered by the expert witness must be based upon a reasonable degree of medical or scientific certainty. That is, in the judgment of the expert witness, the causal relationship between an event and the outcome is probable, or more likely than not. The quality of the **evidence**, therefore, rises above speculation and conjecture and may be considered by the jury.”).
- 157 Lewin, *supra* note 138, at 402-03 & n.88.
- 158 405 F.3d 66, 73 (1st Cir. 2005).
- 159 *Id.* at 90.
- 160 *Id.* at 90-91.
- 161 See *id.* at 91.
- 162 *Id.*; see also *Draper v. United States*, 358 U.S. 307, 313 (1959) (“Probable cause exists where ‘the facts and circumstances within [the arresting officers] knowledge and of which they had reasonably trustworthy information [are] sufficient in themselves to warrant a man of reasonable caution in the belief that’ an offense has been or is being committed.” (quoting *Carroll v. United States*, 267 U.S. 132, 162 (1925))).
- 163 Bradford, *supra* note 152, at 136.
- 164 Lewin, *supra* note 138, at 403-06.
- 165 *Id.* at 402.
- 166 Bradford, *supra* note 152, at 142.
- 167 Lewin, *supra* note 138, at 490-92.
- 168 See, e.g., *State v. Hunt*, 2004 WI App 68, ¶ 3, 271 Wis. 2d 818, 677 N.W.2d 732 (“A DNA test subsequently established to a high degree of medical certainty that Hunt was the father of the baby.”); *United States v. Hammer*, 404 F. Supp. 2d 676, 762 (M.D. Pa. 2005).
- 169 *Howard v. State*, 853 So. 2d 781, 803 (Miss. 2003).
- 170 *Id.*
- 171 *Id.* at 803-04 (“How can Dr. West testify outright that these marks were left by this individual; yet an expert testifying to DNA **evidence** (the most special and unique makeup of our bodies) is not allowed to testify that the blood is the defendant's or the victim's, but rather has to give a statistical probability regarding the likelihood that the blood is the defendant's or victim's? This makes no sense.”).
- 172 See, e.g., Letter from William H. Perloff to Troy Cross, *supra* note 59, at 3; *State v. Ostlund*, 416 N.W.2d 755, 758-59 (Minn. Ct. App. 1988).
- 173 See Bradford, *supra* note 152, at 141.
- 174 See *supra* Part II.C.
- 175 See, e.g., *People v. Wong*, 619 N.E.2d 377, 380 (N.Y. 1993) (“An autopsy performed on the child revealed that he had died as a result of internal brain injuries, including ruptured blood vessels, that could only be attributed to **shaken baby** syndrome.”).
- 176 Telephone Interview with Patrick Turski, Neuroradiologist, Univ. of Wis.-Madison (Nov. 16, 2005); see Barnes, *supra* note 45, at 91.

SHAKEN BABY SYNDROME: MEDICAL UNCERTAINTY..., 2007 Wis. L. Rev. 701

- 177 See Sophia Kazmi, Lawyer Says Caregiver Did Not Injure Baby Girl, *Contra Costa Times* (Walnut Creek, Cal.), Nov. 10, 2005 (“Doctors commonly hear an excuse for injuries, such as the baby fell from a bed or sofa.”); see also Sharp, *supra* note 9; Rachel McCormick, Ex-Raider Charged with Child Abuse, *J. Times* (Racine, Wis.), Oct. 15, 2005, at 13A; Baby Sitter Rejects Plea Deal in **Shaken Baby** Case, *Post-Crescent* (Appleton, Wis.), Dec. 2, 2005, at 1C.
- 178 See, e.g., Sharp, *supra* note 9; *State v. Ostlund*, 416 N.W.2d 755, 758 (1988) (“It is inconceivable that these injuries could have occurred from a fall off a couch.”).
- 179 See, e.g., Plunkett, *supra* note 102, at 10.
- 180 Letter from William H. Perloff to Troy Cross, *supra* note 59, at 3.
- 181 See, e.g., *id.* (referring to witness statements that the child was “playing normally” and then was later “described as abnormal,” and the defendant’s statements defendant that she “possibly” shook her son).
- 182 Interview with Norman Fost, Pediatrician, Univ. of Wis.-Madison, in Madison, Wis. (Nov. 1, 2005).
- 183 See, e.g., Letter from William H. Perloff to Troy Cross, *supra* note 59, at 1.
- 184 See Louis B. Schlesinger, *A Case Study Involving Competency to Stand Trial: Incompetent Defendant, Incompetent Examiner, or “Malingering by Proxy”?*, 9 *Psychol. Pub. Pol’y & L.* 381, 385 (2003). Researchers found that clinical impressions among psychologists were influenced by various patient traits. See *id.* They then would ask specific questions to support their hypothesis. *Id.* It has also been noted that an examiner bias is common in forensic sciences. See *id.*; Peter J. Neufeld, *The (Near) Irrelevance of Daubert to Criminal Justice and Some Suggestions for Reform*, 95 *Am. J. Pub. Health* S107, S111 (Supp. I 2005). An examiner might have been exposed to “irrelevant case information, increasing the likelihood of a false positive.” *Id.*
- 185 See Letter from William H. Perloff to Troy Cross, *supra* note 59, at 1; see also Schlesinger, *supra* note 184, at 385 (“Here, the forensic psychologist finds in the defendant (nonexistent) signs, symptoms, or disorders that were initially suggested by the referring attorney.... The effect, which could be called ‘malingering by proxy,’ derives from the forceful opinions of the legal advocate, which can be quite contagious.”).
- 186 See Geddes et al., *supra* note 108, at 20.
- 187 See Geddes & Plunkett, *supra* note 44, at 719-20.
- 188 See *id.*
- 189 See Barnes, *supra* note 45, at 91.
- 190 See *id.* at 91-92.
- 191 See *Fed. R. Evid.* 702.
- 192 See *Fed. R. Evid.* 702 advisory committee’s note.
- 193 *Commonwealth v. Davis*, No. 04-CR-205, at 21 (Greenup County Cir. Ct. Apr. 17, 2006).
- 194 Le Fanu, *supra* note 65, at 251.
- 195 Sam Lister, Q & A: **Shaken Baby** Syndrome, *Times Online*, Feb. 14, 2006, <http://www.timesonline.co.uk/tol/news/uk/article546383.ece>.
- 196 *R v. Cannings*, [2004] EWCA (Crim) 1, ¶¶ 2, 4, 2 *Crim. App.* 7 (H.L.) (appeal taken from Winchester).
- 197 *Id.* ¶ 22.

SHAKEN BABY SYNDROME: MEDICAL UNCERTAINTY..., 2007 Wis. L. Rev. 701

- 198 Id. ¶ 178.
- 199 R v. Harris & Ors, [2005] EWCA (Crim) 1980 ¶ 4 (Eng.); Lister, supra note 59.
- 200 Lister, supra note 59.
- 201 Harris & Ors, [2005] EWCA (Crim) 1980, at ¶ 1.
- 202 Id. ¶ 3. These four appeals constituted a test case regarding the reliability of medical **evidence** in SBS cases. Lister, supra note 59.
- 203 Harris & Ors, [2005] EWCA (Crim) 1980, at ¶ 153.
- 204 Id. ¶ 13.
- 205 Id. ¶ 152.
- 206 Id. ¶ 44.
- 207 Id. ¶ 45.
- 208 Id. ¶¶ 222-23.
- 209 Id. ¶ 44.
- 210 Id. ¶ 250.
- 211 Id. ¶ 266.
- 212 Lister, supra note 59.
- 213 Harris & Ors, [2005] EWCA (Crim) 1980, at ¶¶ 56-100.
- 214 Id. ¶ 270.
- 215 Id. ¶ 271.
- 216 Id. ¶ 275.
- 217 Lister, supra note 59.
- 218 N. Stoodley, Controversies in Non-Accidental Head Injury in Infants, 79 Brit. J. Radiology 550, 553 (2006).
- 219 Dyer, supra note 43.
- 220 Id.
- 221 Id.
- 222 See id.
- 223 R v. Court [2003] WASCA 308.
- 224 Id. ¶ 9.
- 225 Id. ¶ 11.
- 226 Id.

SHAKEN BABY SYNDROME: MEDICAL UNCERTAINTY..., 2007 Wis. L. Rev. 701

- 227 Id. ¶ 74.
- 228 Id. ¶ 46.
- 229 Id. ¶ 217.
- 230 Guilliatt, *supra* note 33, at 19-22.
- 231 Jane O'Hara, *The Babysitter Didn't Do It*, *Maclean's*, May 14, 2001, at 62.
- 232 See *id.*
- 233 Id. He did not obtain X-rays of the body, speak with the doctor who operated on the baby, or examine the tissue under a microscope before making a diagnosis. Id.
- 234 Id.
- 235 Id. He said “there is simply no doubt. There is only one conclusion I can come to.” Id.
- 236 Id. Her father sold his home and spent over \$150,000 for her defense, which brought in nine experts to refute Smith. Id.
- 237 Id.
- 238 Kirk Makin, *Ontario Orders Pathologist's Work Reviewed*, *Globe & Mail (Can.)*, June 8, 2005, at A8.
- 239 Jane O'Hara, *Dead Wrong*, *Maclean's*, May 14, 2001, at 54, 55.
- 240 Id.
- 241 Makin, *supra* note 238.
- 242 See *id.* O'Hara, *supra* note 231, at 57. Smith later left Ontario and found work in Saskatchewan, but when he applied for his license he did not disclose the Ontario investigation. *Pathologist Reprimanded by Sask. College*, *CBC News*, Feb. 5, 2007, <http://www.cbc.ca/health/story/2007/02/05/smith-reprimand.html>. His Saskatchewan license was then revoked, and he has reapplied. Id.
- 243 See Fung et al., *supra* note 96, at 41.
- 244 Id. at 40.
- 245 Id.
- 246 Id.
- 247 Id.
- 248 Id.
- 249 See *id.* at 40-41.
- 250 Id. at 41. This is obviously problematic because racial and class bias may influence a diagnosis. In the United States, courts may determine that certain defendants do not fit the “profile” for SBS cases. See, e.g., *Smith v. Mitchell*, 437 F.3d 884, 889 (9th Cir. 2006) (“Grandmothers, especially those not serving as the primary care-takers, are not the typical perpetrators.... [T]here was no evidence of any precipitating event that might have caused the Petitioner to snap and assault her grandson. She was not trapped in a hopeless situation with a child she did not want or love. Nor was she forced to single-handedly care for a baby that had been crying all day and all night.”).

SHAKEN BABY SYNDROME: MEDICAL UNCERTAINTY..., 2007 Wis. L. Rev. 701

- 251 Fung et al., supra note 96, at 41.
- 252 Id.
- 253 Eva Lai Wah Fung stated that these case conferences were arranged because of “the western experience.” Id. She mentioned that western doctors who are concerned that child abuse is underdiagnosed question the Japanese study’s description of an “infantile acute subdural hematoma” that can be caused by minor head trauma. Id.
- 254 Id.
- 255 Even a confession is disconcerting because many caregivers might admit to shaking the child when they were frantically trying to revive it after it was not breathing. Lyons, supra note 45, at 1129-30.
- 256 Geddes et al., supra note 108, at 20.
- 257 See id.
- 258 See Ramsey, supra note 41, at 2.
- 259 Guilliat, supra note 33, at 19.
- 260 Id.
- 261 Christopher M. Milroy, Editorial, Medical Experts and the Criminal Courts, 326 BMJ 294, 294 (2003) (“Concerns have been raised where an expert only ever appears for the prosecution or the defence and about experts who seem to be pursuing a sociopolitical agenda not based on objective **evidence**.”).
- 262 [Fed. R. Evid. 703](#). The Advisory Committee notes to the rule emphasize that a physician may rely on data presented outside of the court, including “statements by patients and relatives, reports and opinions from nurses, technicians and other doctors, hospital records, and X-rays.” [Fed. R. Evid. 703](#) advisory committee’s note.
- 263 Id.
- 264 Certain instances may require inquiries about what happened to the baby for treatment purposes, such as how long it was not breathing, or the signs and symptoms observed by the caregiver. See supra Part III.C.
- 265 R v. Harris & Ors [2005] EWCA Crim 1980.
- 266 Cf. [Draper v. United States](#), 358 U.S. 307, 313 (1959).
- 267 There are even prosecutors that are known to be experts on prosecuting SBS cases. See Julie Bykowicz, Father Acquitted in Baby’s Death, Balt. Sun, Feb. 8, 2007, at 1B (“[T]he prosecutor, Assistant State’s Attorney Mary-Ann Burkhart, is a national expert on how to prosecute **shaken-baby** syndrome cases.”).
- 268 Dyer, supra note 43.
- 269 Cf. Stephen Hunt, [Shaken Baby](#) Concept Denied, Salt Lake Trib., July 8, 2004, at B1.
- 270 **Evidence**-based medicine predicates “medical practice and opinions on the best available medical and scientific **evidence**.” Donohoe, supra note 56, at 239.

2007 WILR 701

7 N.Y.U. J. L. & Liberty 221

NYU Journal of Law & Liberty
2013

Article

Paul C. Giannelli^{a1}

Copyright (c) 2013 New York University Journal of Law & Liberty; Paul C. Giannelli

JUNK SCIENCE AND THE EXECUTION OF AN INNOCENT MAN

“The only statement I want to make is that I am an innocent man convicted of a crime I did not commit. I have been persecuted for twelve years for something I did not do.”--Cameron Todd Willingham's words before his execution.¹

I. Introduction

Two-year-old Amber Willingham, along with her younger twin sisters, Karmon and Kameron, died in a fire on December 23, 1991 in Corsicana, Texas. Their father Cameron Todd Willingham escaped from the fire, was tried, and eventually executed for their deaths. The expert testimony offered against him to prove **arson** *222 was “junk science.”² The case has since become infamous-- the subject of an award-winning New Yorker article,³ numerous newspaper accounts,⁴ and several television shows.⁵ It also became enmeshed in the death penalty debate⁶ and the reelection of Texas Governor Rick Perry, who refused to grant a stay of execution after a noted expert submitted a report debunking the **arson** “science” offered at Willingham's trial.⁷ The Governor later attempted to derail *223 an investigation by the Texas Forensic Science Commission into the **arson evidence** presented at Willingham's trial.⁸

Whatever else the Willingham case may stand for, it is a trenchant illustration of the judicial acceptance of expert testimony devoid of empirical support and the legal system's inability to effectively police such testimony. The National Academy of Science's landmark 2009 report on forensic science, *Strengthening Forensic Science in the United States: A Path Forward*, made the breathtaking observation that, “[a]mong existing forensic methods, only nuclear DNA analysis has been rigorously shown to have the capacity to consistently, and with a high degree of certainty, demonstrate a connection between an evidentiary sample and a specific individual or source.”⁹ The report went on to observe: “In a number of forensic science disciplines, forensic science professionals have yet to establish either the validity of their approach or the accuracy of their conclusions, and the courts have been utterly ineffective in addressing this problem.”¹⁰ Moreover, recent studies document the role that forensic science played in convicting the innocent.¹¹ The *224 Willingham case also highlights the corrosive effect of death-penalty politics--the extraordinary lengths a state has undertaken to avoid investigating the possibility that it had executed an innocent man.

II. The Trial

A. The **Arson Evidence**

JUNK SCIENCE AND THE EXECUTION OF AN..., 7 N.Y.U. J. L. &...

Willingham's capital trial began in August 1992. Proclaiming his innocence from the beginning, Willingham refused to plead guilty in exchange for life imprisonment in lieu of the death penalty.¹²

***225** 1. **Arson** Investigations

The **arson evidence** was critical. No **arson**, no crime.¹³ The prosecution proffered two experts: Manuel Vasquez, a deputy state fire marshal, and Douglas Fogg, an assistant fire chief in Corsicana.¹⁴ With years of experience, they came from the “old school” of investigators--those who used intuition and a number of rules of thumb to determine whether a fire was incendiary.¹⁵ In Vasquez's words: “The fire tells a story. I am just the interpreter. . . . And the fire does not lie. It tells me the truth.”¹⁶

Critics of this approach complained that it lacked a scientific foundation. Rather, it was based on folklore that had been passed down from generation to generation without any empirical testing.¹⁷ As early as 1977, a government report noted that common **arson** indicators had “received little or no scientific testing” and “[t]here appears to be no published material in the scientific literature to substantiate their validity.”¹⁸ Through the 1980s, proponents of a science-based approach to **arson** investigations waged an uphill battle, finally winning a major victory in 1992 when the National ***226** Fire Protection Association (NFPA) published its Guide for Fire and Explosion Investigations (NFPA 921).¹⁹

Although NFPA 921 would subsequently become the bible in fire and **arson** investigations,²⁰ it was published weeks after Willingham's trial.

2. The Willingham Fire

Deputy Fire Marshal Vasquez told the jury that he had found twenty indicators of **arson** during his post-fire investigation of Willingham's house.²¹ One indicator was a low burning fire.²² “All fire goes up,” Vasquez testified.²³ Thus, burn patterns on the lower walls and floor suggested that an accelerant was used.²⁴

***227** This common-sense notion, however, has its limitations, especially when a fire occurs in a contained area, such as a house. Due to buoyancy, a thermal plume initially rises once a fire is ignited. As the fire continues, the plume reaches the ceiling, which causes it to spread outward towards the walls. When it reaches the walls, the combustion products press down from the ceiling creating an upper level, which continues to increase in depth and temperature. Eventually thermal radiation replaces convection as the principal method of heat transfer. When the temperature of the hot gas layer reaches approximately 1100-1200 degrees Fahrenheit, every exposed combustible surface in the room will burst into flames. This phenomenon, known as “flashover,” can occur within minutes. After flashover, the entire room is engulfed in flames, including the lower walls and floor.²⁵ Flashover, according to one authority, is the point at which the fire transitions from a “fire in a room” to a “room on fire.”²⁶ Consequently, a low burning fire is not necessarily indicative of an incendiary origin.

Moreover, some of Vasquez's other “indicators”--splotchy looking areas called “puddle configurations” and “pour patterns”--are present after flashover in an accidental fire.²⁷ Similarly, additional indicators such as “alligatoring” marks on wood (char blisters) are explained by flashover. Flashover also accounts for another fact that Vasquez thought incriminatory. Willingham told investigators ***228** that he had attempted to save his daughters, but the heat was too great and he was forced to run from the house without shoes. Willingham's feet were not burnt, and in Vasquez's mind, burnt debris on the floor made that impossible.²⁸ However, if Willingham left his home before flashover, his feet would

JUNK SCIENCE AND THE EXECUTION OF AN..., 7 N.Y.U. J. L. &...

not have been burnt. (Willingham exaggerated his attempts to save his children--a common occurrence when a parent survives a fatal fire.)

Still another clue was charring under an aluminum threshold of an interior door. Here, again, this may occur in a flashover. Still other **arson** indicators--melted bedsprings,²⁹ multiple points of origin,³⁰ and brown stains on a concrete floor³¹--are also consistent with an accidental blaze.³² In addition, Vasquez relied on the presence of "crazed glass"--spider-web patterns on the windows as an ***229** indication of **arson**.³³ It was long believed that crazed glass resulted from a fire that burned fast and hot--i.e., one fueled by a liquid accelerant. Yet, subsequent research demonstrated that crazing occurs only from rapid cooling when water from fire hoses is sprayed on heated windows.³⁴

In retrospect, the most damning piece of **evidence** involved one of the numerous debris samples submitted for laboratory analysis.³⁵ It came from an area near the front door and was the only sample that tested positive for a chemical commonly used in charcoal lighter fluids. Nevertheless, this finding can be explained by the fact that a charcoal grill and lighter fluid were on the front porch at the time of the fire.³⁶ Eyewitnesses reported no fire at the front door when they first saw Willingham on the porch. In fact, the negative results from the other samples support Willingham's case.³⁷

***230** The cause of the fire remains unknown,³⁸ and the scene cannot be reconstructed due in part to the disappearance of records.³⁹

3. Credibility Testimony

Vasquez did not limit himself to an opinion on the cause of the fire. He also testified that Willingham was not truthful, informing the jury that during an interview Willingham "told me a story of pure fabrication"⁴⁰ and, "[h]e just talked and he talked and all he did was lie."⁴¹ This testimony was improper and extremely prejudicial. Vasquez was accepted by the court as an **arson** investigator, not as an expert on credibility. He was thus testifying beyond his ***231** expertise.⁴² Indeed, it is axiomatic that witnesses, lay and expert, are not permitted to testify about credibility.⁴³ Moreover, research suggests that police and other investigators are not all that good at judging credibility: "Unfortunately, psychological research has generally failed to support the claim that individuals [such as CIA, FBI, and police investigators] can attain high levels of performance in making judgments of truth and deception. Over the years, numerous studies have demonstrated that these individuals perform at no better than chance level in detecting deception."⁴⁴

***232** Remarkably, Vasquez also testified that Willingham's "intent was to kill the little girls."⁴⁵ Here, again, the testimony was far beyond his expertise. A qualified **arson** investigator may be able to determine whether a fire was intentionally set but not why it was set--i.e., whether it was set for insurance money, vengeance, etc. Other parts of Vasquez's testimony were also beyond the scope of a fire investigator's expertise. For example, he claimed that Willingham's injuries, including singed eyelids and hair, were self-inflicted.⁴⁶

B. Jailhouse Informant

Johnny Webb, a jailhouse informant, was another prosecution witness.⁴⁷ For obvious reasons, jailhouse snitches are notoriously unreliable. As Judge Trott, a former prosecutor, has observed, "[t]he most dangerous informer of all is the

jailhouse snitch who claims another prisoner has confessed to him.”⁴⁸ According to the Innocence *233 Project, such testimony appeared in eighteen percent of the cases in which convicts were subsequently exonerated by DNA profiling.⁴⁹

Like many informants, Webb was a drug addict (crack cocaine) who had a criminal record (aggravated robbery, car theft, selling marijuana, theft, and forgery).⁵⁰ He also suffered from post-traumatic stress disorder as a result of a prison rape. Indeed, during cross-examination Webb claimed that he could not remember the crime for which he pled guilty (aggravated robbery): “I could have done it, but I do not remember doing it.”⁵¹ Webb, who was serving a fifteen-year sentence, testified that no promises had been made to him, which in itself is suspect. Implied inducements to informants are well known in criminal practice.⁵² Five years later the prosecutor *234 asked the Texas Board of Pardons and Paroles to grant Webb parole.⁵³

Moreover, Webb's assertions were inherently problematic. He was not Willingham's cellmate. Instead, Webb claimed Willingham told him, a virtual stranger, of his misdeeds through a hole in a steel door in Willingham's cell. Yet, Willingham went to his grave proclaiming his innocence. Webb also asserted that Willingham said he started the fire to hide his wife's abuse of their children;⁵⁴ there was no **evidence** that Willingham or his wife, Stacy, ever abused their children. Later, Webb recanted his testimony and then retracted the recantation. A journalist would later recount an interview with Webb: “After I pressed him, [Webb] said, ‘It's very possible I misunderstood *235 what [Willingham] said.’ . . . He paused, then said, ‘The statute of limitations has run out on perjury, hasn't it?’”⁵⁵

C. Demeanor **Evidence**

The other type of **evidence** involved Willingham's behavior--that is, testimony that he made insufficient efforts to save his children and did not show sufficient grief at the hospital or the next day.⁵⁶ The prosecution emphasized this in his final argument.⁵⁷ Not surprisingly, the **evidence** regarding the fire scene is somewhat conflicting. Several neighbors, who testified for the prosecution, acknowledged that Willingham “was hollering. He was screaming the babies was in there.”⁵⁸ A paramedic testified that Willingham *236 was “really excited” and “hysterical.”⁵⁹ A police officer stated that Willingham was “upset” and “[w]e had to end up restraining him a little bit.”⁶⁰ There was also testimony that Willingham had attempted to reenter the house by breaking several windows.⁶¹

The reactions of persons to traumatic events are far too varied to place much weight on their demeanor, and this includes survivors of fires.⁶² Further, this **evidence** changed over time, becoming more damaging after the investigators became convinced that Willingham was an arsonist.⁶³ Once witnesses learn of investigators' suspicions, it is not unusual for their testimony to harden and become *237 more definitive.⁶⁴ Moreover, similar “demeanor” **evidence** has proved unreliable in other **arson** cases.⁶⁵

D. Motive **Evidence**

“[T]here was no clear motive. The children had life insurance policies, but they amounted to only fifteen thousand dollars, and Stacy's grandfather, who had paid for them, was listed as the primary beneficiary.”⁶⁶ Moreover, neither Willingham nor his wife knew of the insurance until after the fire.⁶⁷

JUNK SCIENCE AND THE EXECUTION OF AN..., 7 N.Y.U. J. L. &...

The only prosecution **evidence** concerning motive is found in the jailhouse informant's testimony. Recall that Johnny Webb testified that Willingham had told him the fire was started to hide Willingham's wife's abuse of their children: "one of the babies were injured or dead or something like that."⁶⁸ There was no **evidence** in the record that either Willingham or his wife had ever abused their children, and the medical **evidence** concerning the autopsies did not support such a claim.⁶⁹ (At the time, this motive may have made ***238** sense to the prosecutor because Willingham's wife Stacy supported him at trial.⁷⁰) The prosecutor did not refer to this **evidence** in his closing statement. Instead, he demonized Willingham with the demeanor **evidence**.

E. Defense Case

Willingham did not take the stand. Apparently, he wanted to testify, but his lawyers thought he would not make a good witness.⁷¹ Willingham's baby sitter, a defense witness, testified that Willingham would not hurt his children. Another defense witness, an incarcerated felon, was proffered in an attempt to impeach Webb, but his testimony was ruled hearsay.

F. Guilt Phase

In sum, the demeanor **evidence** was not very probative,⁷² and the snitch testimony was not reliable. The key to the conviction (and arrest) was the expert testimony. No **arson**, no crime, no arrest.⁷³

***239 III. Pardon & Clemency Proceedings**

Willingham lost his appeal to the Texas Court of Criminal Appeals in 1995.⁷⁴ When other attempts at judicial redress also failed,⁷⁵ his execution date was set for February 17, 2004. At this point, his only hope was clemency, a process that is initiated in the Board of Pardons and Paroles before an application goes to the governor. By this time Willingham's appellate attorney had contacted Dr. Gerald Hurst, a nationally recognized **arson** expert with a chemistry degree from Cambridge University. Working pro bono, Hurst reviewed the **evidence** (e.g., the fire marshal report, trial testimony, photographs, and a 52-minute video of the scene) and prepared a report, concluding that the **arson** testimony was invalid:

***240** The fire investigation report of the Texas State Fire Marshal's Office in this case is a remarkable document. On first reading, a contemporary fire origin and cause analyst might well wonder how anyone could make so many critical errors in interpreting the **evidence**. However, when the report is looked at in the context of its time and in light of a few key advances that have been made in the fire investigation field in the last dozen years, it becomes obvious that the report more or less simply reflects the shortcomings in the state of the art prior to the beginning of serious efforts to introduce standards and to test old theories that had previously been accepted on faith.⁷⁶

The five page report, which methodically examined the major deficiencies of the Willingham fire investigation, was submitted four days before the execution to the pardon board and to Governor Rick Perry. Notwithstanding this report, which raised substantial questions about the origin of the fire, the state of Texas executed Willingham by lethal injection as scheduled.⁷⁷ Whether either the board or the Governor ever read the report is unclear.⁷⁸

*241 IV. Texas Forensic Science Commission

After Willingham's execution, two seemingly unrelated statutes were enacted that ensured that the case would not die. In November 2004, Congress passed the Justice for All Act.⁷⁹ Because of numerous crime laboratory scandals,⁸⁰ this legislation included a requirement that each state receiving federal funds designate an entity to investigate forensic misconduct and incompetence.

One of the major scandals involved the Houston crime laboratory.⁸¹ According to a state senator, “the validity of almost any case that has relied upon **evidence** produced by the lab is questionable.”⁸² As described by a later investigation, “the DNA Section was *242 in shambles--plagued by a leaky roof, operating for years without a line supervisor, overseen by a technical leader who had no personal experience performing DNA analysis and who was lacking the qualifications required under the FBI standards, staffed by underpaid and undertrained analysts, and generating mistake-ridden and poorly documented casework.”⁸³ As a consequence, the state legislature created the Texas Forensic Science Commission (TFSC) in 2005.⁸⁴ Among other duties, the Commission was tasked with investigating claims of professional negligence or misconduct as required by the federal act.

A. Innocence Project Complaint

By this time, the Chicago Tribune, after reviewing the Hurst report, began examining the Willingham case.⁸⁵ The Tribune retained three independent experts to review the **arson evidence**, all of whom concluded that the **evidence** was seriously flawed.⁸⁶ Next, the Innocence Project requested five experts to reexamine the case pro bono. These experts submitted a scathing forty-three page report, finding that “each and every one of the indicators relied upon *243 have since been scientifically proven to be invalid.”⁸⁷ The report even raised questions about Fire Marshal Vasquez's general knowledge of the field. For example, Vasquez testified that of the 1200 to 1500 fires he had investigated, most were **arson**.⁸⁸ Yet, the Texas Fire Marshal Office reported that between 1980 and 2005, only fifty percent of investigated fires were **arson**.⁸⁹ Vasquez also testified that fifty percent of his fires involved injuries and deaths.⁹⁰ In contrast, between 1995 and 2005, the annual percentage of fires that resulted in death was 0.23% and the percentage of those resulting in injuries was 1.22%.⁹¹

In May 2006, the Innocence Project petitioned the Commission to review the **arson** testimony in the Willingham and Ernest Ray Willis cases.⁹² The expert **evidence** in both cases was comparable, but Willis was lucky. His death penalty conviction was overturned on procedural grounds,⁹³ and the prosecutor subsequently refused to reindict him after Dr. Hurst wrote the same type of critical report in Willis's case that he had written in Willingham's.⁹⁴ Willis, who *244 had spent seventeen years on death row, was subsequently exonerated on grounds of actual innocence.⁹⁵

The TFSC was not authorized to determine guilt or innocence. Instead, the Innocence Project argued that the State Fire Marshal's Office should have reinvestigated the Willingham and other old **arson** cases, in which its experts had testified, after NFPA 921 was published in 1992--a full twelve years before Willingham's execution.⁹⁶

B. The Beyler Report

The Commission's work was hampered from the beginning. Initially, the legislature did not provide funding, and then the Governor and Lieutenant Governor delayed the appointment of Commission members.⁹⁷ When funding was finally appropriated, the Commission spent a year formulating its procedures under the guidance of the Texas Attorney General's Office. In late 2008, more than two years after the Innocence Project complaint was received, the Commission retained an independent consultant, Dr. Craig Beyler, another nationally recognized expert, to review the **arson evidence**.⁹⁸ *245 Beyler's fifty-one page report dissected the expert testimony, concluding:

The investigations of the Willis and Willingham fires did not comport with either the modern standard of care expressed by NFPA 921, or the standard of care expressed by fire investigation texts and papers in the period 1980-1992. The investigators had poor understandings of fire science and failed to acknowledge or apply the contemporaneous understanding of the limitations of fire indicators. Their methodologies did not comport with the scientific method or the process of elimination. A finding of **arson** could not be sustained based upon the standard of care expressed by NFPA, or the standard of care expressed by fire investigation texts and papers in the period 1980-1992.⁹⁹

C. Governor's Intervention

Once Beyler's report became public, a political firestorm erupted and Governor Perry, who was in the midst of a reelection battle, replaced Commission members two days before a scheduled hearing to consider Dr. Beyler's report.¹⁰⁰ The newly appointed chair, *246 John Bradley, a prosecutor, cancelled the meeting.¹⁰¹ The timing of the Governor's action raised the specter of a cover-up.¹⁰² Bradley then raised procedural objections, arguing for closed-door meetings, training, development of written policies, and definitions of the terms "negligence" and "misconduct."¹⁰³ Bradley next prepared a report exculpating the Willingham fire investigators of any negligence. The other Commission members, most of whom were scientists, balked,¹⁰⁴ thwarting Bradley's "attempt to turn the science commission into a legalistic briar patch."¹⁰⁵ Governor Perry responded *247 by saying that "'the **evidence** shows Willingham to be guilty,' and [dismissing] the work of Beyler and other **arson** experts. The Governor declined to specify . . . what **evidence** he believes backs up the case."¹⁰⁶ Before the Commission could reconvene, Governor Perry was reelected. By this time, the Fire Marshal's Office and the city of Corsicana were challenging the TFSC's jurisdiction to review old cases.¹⁰⁷

At its January 7, 2011 meeting, the Commission finally heard from Dr. Beyler, who once again reiterated his position that the Willingham investigation was seriously flawed and the cause of the fire should have been designated as "undetermined."¹⁰⁸ In his view, the investigators failed to eliminate natural or accidental causes, in violation of professional standards.¹⁰⁹ Another **arson** expert, Dr. John DeHaan, author of a standard text in the field,¹¹⁰ agreed with Beyler. According to DeHaan, "'everything that was documented *248 post-fire was consistent with accidental rather than intentional fire. There was no basis for concluding that this was **arson**.' DeHaan said he was dismayed that the state fire marshal's office stood by the conclusion of the investigators."¹¹¹ In contrast, Ed Salazar, a lawyer with the State Fire Marshal Office, defended the fire investigation. One report put it this way: "Salazar became impassioned with his criticism of the opposing expert, but he was short on analysis. 'It was embarrassing,' said one scientist on the commission afterward."¹¹² Buddy Wood, a senior investigator with the Houston Fire Department, testified that the investigators were not negligent. However, he also stated that he could not determine the cause of the fire because he had not gone to the scene.¹¹³

On January 28, 2011 the Commission requested a legal opinion regarding its jurisdiction from the state Attorney General's Office.¹¹⁴

D. TFSC Report

While awaiting the Attorney General's response, the Commission issued a limited report--one that did not directly deal with the Willingham and Willis cases.¹¹⁵ Nevertheless, the report's recommendations *249 and statements indicated that the Willingham **arson** investigation was seriously flawed. Its first recommendation was "that fire investigators adhere to the standards of NFPA 921."¹¹⁶ In addition, the report reviewed a number of the **arson** indicators that were used in the Willingham and Willis cases. Reviewing Vasquez's testimony, the report undermined his opinions concerning V-patterns as an indicator of origin, pour patterns, low/deep burning, multiple separate points of origin, spalling, burn intensity, and crazed glass.¹¹⁷ It also observed that, "testimony, such as Vasquez's response to a question regarding Willingham's state of mind, is an example of the type of testimony that experts should avoid as falling outside of their field of expertise."¹¹⁸ Surprisingly, the report even encouraged lawyers to "aggressively pursue admissibility hearings in **arson** cases."¹¹⁹

The State Fire Marshal's Office was criticized as well. That office had submitted a letter that included the following statement: "In reviewing documents and standards in place then and now, we *250 stand by the original investigator's report and conclusions."¹²⁰ In response, the report commented: "This appears to be an untenable position in light of advances in fire science. The fires in these cases occurred two decades ago; there are few circumstances in which an investigation could not be improved with the benefit of twenty years of controlled scientific experiment and practical experience."¹²¹

Significantly, the report also pointed out that forensic disciplines have a "(1) duty to correct; (2) duty to inform; (3) duty to be transparent; and (4) [[duty to] implement] corrective action" when new scientific knowledge develops.¹²²

V. Conclusion

There is little dispute that the **arson evidence** in Willingham's case, based on myths that had permeated fire investigations for years, was invalid. Every independent expert, including the top experts in the country, has concluded that there was no **evidence** of **arson**.¹²³ Without the **arson evidence**, there never would have been *251 an arrest, much less a trial or execution. The other **evidence** introduced at trial (e.g., jail-informant testimony and demeanor **evidence**) was suspect.¹²⁴ While Willingham was executed, Willis, who was convicted on comparable **evidence**, was exonerated after spending seventeen years on death row.

Although NFPA 921 was published in 1992 just weeks after Willingham's trial, many of its findings had been reported during the prior decade. After NFPA 921 was published, the State Fire Marshal's Office became aware of its contents¹²⁵ but did not take corrective action in old cases during the dozen years Willingham waited on death row. Indeed, the Office still maintains it did everything right.

Moreover, once Dr. Hurst's report was made available to them, the Texas pardon board and Governor Perry had the opportunity to stay the execution to investigate further. They did not. Although the United States Supreme Court has

JUNK SCIENCE AND THE EXECUTION OF AN..., 7 N.Y.U. J. L. &...

called clemency the “failsafe” *252 procedure in death penalty cases,¹²⁶ the Texas procedure was known as “death by fax” because the pardon board is not required to meet or discuss a case;¹²⁷ each member can vote by fax.¹²⁸ “Between 1976 and 2004, when Willingham filed his petition, the State of Texas had approved only one application for clemency from a person on death row.”¹²⁹ In another case, a Texas appellate judge wrote: “Applicant's complaints about the inadequacies of our Texas executive clemency procedures are not unheard of. Not only are they not unheard of, but her complaints are pretty much accurate. I would say that clemency law in Texas is a legal fiction at best.”¹³⁰

*253 Then, by interfering with the work of the Texas Forensic Science Commission, Governor Perry and his allies undermined a process intended to improve expert testimony in criminal prosecutions. Congress enacted the requirement that each state designate an entity to investigate forensic misconduct and incompetence because few states had such a procedure and the experience in the states varied.¹³¹ Thus, the TFSC's decision to review the Willingham case was historic--one of the first investigations by a forensic commission in the country. Unfortunately, the Fire Marshal's Office's resistance to admitting prior mistakes and the Governor's intervention undermined the Commission's work. There are still inmates in Texas prisons who were convicted on the same flawed **arson evidence**.¹³²

Footnotes

^{a1} Distinguished University Professor & Weatherhead Professor, Case Western Reserve University; J.D., 1970, LL.M. 1975, University of Virginia; M.S. Forensic Science, 1973, George Washington University.

¹ See David Grann, Trial by Fire: Did Texas Execute an Innocent Man?, *The New Yorker*, Sept. 7, 2009, at 63.

² Christy Hoppe, Some Experts Question Science in Texas **Arson** Cases, *Charleston Gazette & Daily Mail* (W. Va.), Sept. 20, 2009, at 11A (“**Arson** investigators in Texas have relied on old wives' tales and junk science to send men to prison, and perhaps even the death chamber, top experts on fire behavior say.”).

³ See Grann, *supra* note 1. Grann's article won the 2009 George Polk Award for Magazine Reporting, the America Bar Association's 2010 Silver Gavel Award for Media and the Arts, and the 2009 Sigma Delta Chi Award for magazine investigative reporting from the Society of Professional Journalists.

⁴ See, e.g., Jennifer Emily, Texas Forensic Science Commission Refuses to End Inquiry into Willingham **Arson** Case, *Dall. Morning News*, Sept. 18, 2010; Steve Mills & Maurice Possley, Texas Man Executed on Disproved Forensics: Fire that Killed his 3 Children Could Have Been Accidental, *Chi. Trib.*, Dec. 9, 2004, at C1; Steve Mills, Texas May Have Put Innocent Man to Death, Panel Told; Nobody Would Listen, Lawyer, Expert Say, *Chi. Trib.*, Apr. 20, 2005, available at http://articles.chicagotribune.com/2005-04-20/news/0504200146_1_cameron-todd-willingham-willingham-case-gerald-hurst.

⁵ See *Frontline: Death by Fire* (PBS television broadcast Oct. 19, 2010); *Nightline* (ABC television broadcast Sept. 17, 2009). A documentary film, *Incendiary*, has also been released.

⁶ See Emily, *supra* note 4 (“The Willingham case has drawn worldwide attention from death-penalty opponents and others since questions were raised about the integrity of the science **evidence** used to convict him of murder.”). Justice Scalia once wrote: “It should be noted at the outset that the dissent does not discuss a single case--not one--in which it is clear that a person was executed for a crime he did not commit. If such an event had occurred in recent years, we would not have to hunt for it; the innocent's name would be shouted from the rooftops by the abolition lobby.” *Kansas v. Marsh*, 548 U.S. 163, 188 (2006).

⁷ See Mary Alice Robbins, Fired Up: Changes Sought for Texas Forensic Science Commission at Center of Heated Controversy, 25 *Tex. Lawyer*, Nov. 9, 2009 (“Anti-death penalty activists have contended that Willingham was innocent and that [Governor]

JUNK SCIENCE AND THE EXECUTION OF AN..., 7 N.Y.U. J. L. &...

Perry replaced the commission members to block a review of a report questioning whether the fire Willingham was accused of starting was **arson**.”).

8 See infra text accompanying notes 100-15.

9 Nat'l Research Council, National Academy of Sciences, Strengthening Forensic Science in the United States: A Path Forward 100 (2009).

10 Id. at 53 (emphasis added). The Report devotes only two paragraphs to **arson** investigations: “Despite the paucity of research, some **arson** investigators continue to make determinations about whether or not a particular fire was set. However, according to testimony presented to the committee, many of the rules of thumb that are typically assumed to indicate that an accelerant was used (e.g., ‘alligatoring’ of wood, specific char patterns) have been shown not to be true. Experiments should be designed to put **arson** investigations on a more solid scientific footing.” Id. at 173 (emphasis added).

11 A study of 200 DNA exonerations found that forensic **evidence** (fifty-five percent) was the second leading type of **evidence** (after eyewitness identifications, seventy-nine percent) used in the **wrongful conviction** cases. Of the types of forensic **evidence** introduced at trial, “serological analysis of blood or semen [was] the most common (79 cases), followed by expert comparison of hair **evidence** (43 cases), soil comparison (5 cases), DNA tests (3 cases), bite mark **evidence** (3 cases), fingerprint **evidence** (2 cases), dog scent (2 cases), spectrographic voice **evidence** (1 case), shoe prints (1 case) and fiber comparison (1 case).” Brandon L. Garrett, *Judging Innocence*, 108 *Colum. L. Rev.* 55, 81 (2008). See also Brandon L. Garrett & Peter J. Neufeld, *Invalid Forensic Science Testimony and Wrongful Convictions*, 95 *Va. L. Rev.* 1, 14-15 (2009) (“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 which transcripts were located, 25 cases, or 38%, had invalid forensic science testimony.”); Paul C. Giannelli, *Wrongful Convictions and Forensic Science: The Need to Regulate Crime Labs*, 86 *N.C. L. Rev.* 163 (2007) (discussing lab scandals in West Virginia, Oklahoma City, Chicago, Houston, Virginia, Montana, and the FBI Lab).

12 Willingham's court-appointed trial attorneys, John Martin and Robert Dunn, advised him to accept the offer, but he refused. “Willingham was implacable. ‘I ain't gonna plead to something I didn't do, especially killing my own kids,’ he said. It was his final decision. Martin says, ‘I thought it was nuts at the time--and I think it's nuts now.’” Grann, *supra* note 1, at 48. “Though his father did not believe that he should plead guilty if he were innocent, his stepmother beseeched him to take the deal. ‘I just wanted to keep my boy alive,’ she told me.” Id.

13 See generally Paul C. Giannelli & Edward J. Imwinkelried, *Scientific Evidence* ch. 26 (4th ed. 2007) (discussing legal and scientific issues associated with **arson evidence**).

14 Under state law, the Texas State Fire Marshal is responsible for investigating suspicious fires. *Tex. Gov't Code* § 417.007 (West 2004).

15 “Often, the bulk of an investigator's training came on the job, learning from ‘old-timers’ in the field, who passed down a body of wisdom about the telltale signs of **arson**” Grann, *supra* note 1, at 58.

16 Transcript of Statement of Facts, *State v. Willingham*, (Tex.13th D. 1991) (No. 24240-CR) (vol. XI), at 244 [hereinafter Willingham Transcript], *aff'd*, *Willingham v. State*, 897 S.W.2d 351, 354 (Tex. Crim. App. 1995).

17 See John J. Lentini, *Scientific Protocols for Fire Investigation* ch. 8 (2006) (discussing myths of **arson** investigations).

18 J.F. Boudreau et al., National Institute of Law Enforcement and Criminal Justice, Law Enforcement Assistance Administration, U.S. Dep't of Justice, *Arson and Arson Investigation: A Survey and Assessment* 88 (1977).

19 National Fire Protection Association, *Guide for Fire and Explosion Investigation* (1992). The NFPA promotes fire prevention and safety. The most recent edition of NFPA 921 was published in 2011.

20 See *United States v. Hebshie*, 754 F. Supp. 2d 89, 109 n.39 (D. Mass. 2010) (NFPA 921 “is widely accepted as the standard guide in the field of fire investigation.”); Thomas M. May, *Fire Pattern Analysis, Junk Science, Old Wives Tales, and Ipse*

JUNK SCIENCE AND THE EXECUTION OF AN..., 7 N.Y.U. J. L. &...

Dixit: Emerging Forensic 3D Imaging Technologies to the Rescue?, 16 Richmond J.L. & Tech. 1, 5 (2010) (noting that NFPA 921 has “become the de facto national standard for fire scene examination and analysis”).

21 Assistant Chief Fogg's testimony essentially tracked Vasquez's.

22 Vasquez testified that there was “char burning, like, for example, this is the bottom here. It's burned down here at the bottom. That is an indicator in my investigation of an origin of fire because it's the lowest part of the fire.” Willingham transcript, supra note 16, vol. XI, at 239. See also [Willingham v. State, 897 S.W.2d 351, 354 \(Tex. Crim. App. 1995\)](#) (“An expert witness for the State testified that the floors, front threshold, and front concrete porch were burned, which only occurs when an accelerant has been used to purposely burn these areas. This witness further testified that this igniting of the floors and thresholds is typically employed to impede firemen in their rescue attempts.”).

23 Willingham transcript, supra note 16, vol. XI, at 232.

24 “So when I found that the floor is hotter than the ceiling, that's backwards, upside down. It shouldn't be like that. The only reason that the floor is hotter is because there was an accelerant.” Id. at 256.

25 See id. at 75 (“The windows, the electricity started crackling and popping, and the top of the well--well, I was facing the side of the house, and it just blew out. The flames just blew out.... All the windows and the front room was engulfed.”) (testimony of Dianne Barbe); id. at 96 (“We was running towards the house, me and my mother, we was fixing to go and try to get in, and that's when it was an explosion”) (testimony of Brandy Barbe). Vasquez mentioned flashover in his testimony (See id. vol. XII. at 47-48), but he does not appear to understand its implications.

26 Lentini, supra note 17, at 68-70.

27 According to Vasquez, a burn “trailer” was etched on the floor. Willingham Transcript, supra note 16, vol. XI, at 244 (“You can see that on the burnt patterns on this puddle configuration on Exhibit No. 36. This is a strong indicator of a liquid.”).

28 “There was fire on the floor.... He had no injuries on his feet.” Id. at 267.

29 “[T]he springs were burned from underneath. This indicates there was a fire under this bed because of the burn underneath the bed.” Id. at 241.

30 “Multiple areas of origin indicate--especially if there is no connecting path, that they were intentionally set by human hands.” Id. at 255. There are two problems here. First, the fire scene did not exhibit multiple origins, according to independent experts. See Douglas Carpenter et al., Report on the Peer Review of the Expert Testimony in the Cases of State of Texas v. Cameron Todd Willingham and State of Texas v. Earnest Ray Willis 11-12 (2006). Second, even if the fire scene had shown multiple points of origin, this would not necessarily indicate an intentional fire. Lentini, supra note 17, at 461-62.

31 Willingham Transcript, supra note 16, vol. XI, at 248-49. Fire experts reviewing the **evidence** from Willingham's trial pointed out that “[t]he behavior of concrete in fires, including the development of various colors, has been extensively studied.” Carpenter et al., supra note 30, at 18. These experts concluded that there is simply “no scientific basis for Mr. Vasquez's statement about the brown discoloration being an indication of the presence of accelerants.” Id.

32 Vasquez's testimony also demonstrated other misconceptions. A common one is that **arson** fires burn hotter and faster than “normal” fires: “You know, it makes the fire hotter. It's not a normal fire.” Willingham Transcript, supra note 16, vol. XI, at 249. However, the temperature of burning wood and burning gasoline are nearly identical, so to claim that a fire using liquid accelerants burns “hotter” than a wood fire is wrong. Lentini, supra note 17, at 465.

33 “The pieces of broken window glass on the ledge of the north windows to the northeast bedroom disclosed a crazed ‘spider webbing’ condition. This condition is an indication that the fire burned fast and hot.” Carpenter et al., supra note 30, at 18 (citing Vasquez's written report on the Willingham fire).

34 Lentini, supra note 17, at 439 (“It is unclear why anyone ever thought that crazing of glass indicated rapid heating.”).

JUNK SCIENCE AND THE EXECUTION OF AN..., 7 N.Y.U. J. L. &...

- 35 In closing argument, the defense counsel referred to a “dozen samples.” Willingham Transcript, supra note 16, vol. XIII, at 20.
- 36 Id., vol. XII, at 15 (although photographs show a grill, Vasquez apparently did not know of the grill's presence); id. at 16 (acknowledging that a fire-damaged charcoal lighter fluid container was found on the front porch).
- 37 The prosecutor would later say that he “‘never did understand why they weren't able to recover’ positive tests in these parts.” Grann, supra note 1, at 61. At trial, he argued that the “liquid burned away in that destructive madness created by Cameron Todd Willingham.” Willingham Transcript, supra note 16, vol. XIII, at 45.
- 38 Willingham's defense suggested at trial that Amber, the two-year old, could have caused the fire by knocking over a kerosene lamp. This seems like speculation. In his 2009 New Yorker article, Grann raised another theory-- a space heater in the children's bed room--which was never confirmed. Both Willingham and his wife, Stacy, had warned Amber not to play with it. Willingham told investigators that “[h]e and Stacy used three space heaters to keep the house warm, and one of them was in the children's room. ‘I taught Amber not to play with it,’ he said, adding that she got ‘whuppings every once in a while for messing with it.’” Grann, supra note 1, at 46. Although Vasquez testified that the heater was off when he inspected the premises four days after the fire, Stacy said it was on when she left the house on the morning of the fire. Elizabeth Gilbert, who befriended Willingham in 2000 when he was on death row, eventually began to investigate his case after initially believing that he was guilty. During this process, she conducted a taped interview with Stacy: “Stacy was sure that, at least on the day of the incident--a cool winter morning--it had been on. ‘I remember turning it down,’ she recalled. ‘I always thought, Gosh, could Amber have put something there?’ Stacy added that, more than once, she had caught Amber ‘putting things too close to it.’” Id. at 51.
- 39 Report of the Texas Forensic Science Comm'n, Willingham/Willis Investigation 21 (April 15, 2011) (hereinafter TFSC Report) (“Although the [[Corsicana Fire Department] informed the Commission that a thorough examination was conducted, the documentation provided to the District Attorney no longer exists.”).
- 40 Willingham Transcript supra note 16, vol. XI, at 258.
- 41 Id. at 260.
- 42 In addition to a lack of expertise, some courts cite the jury's historic role as a reason for rejecting opinions on credibility. In rejecting polygraph **evidence**, for example, Justice Thomas wrote of the importance of “[p] reserving the court members' core function of making credibility determinations in criminal trials. A fundamental premise of our criminal trial system is that ‘the jury is the lie detector.’” *United States v. Scheffer*, 523 U.S. 303, 312-13 (1998) (quoting *United States v. Barnard*, 490 F.2d 907, 912 (9th Cir. 1973)). Justice Thomas further stated: “Determining the weight and credibility of witness testimony, therefore, has long been held to be the ‘part of every case [that] belongs to the jury, who are presumed to be fitted for it by their natural intelligence and their practical knowledge of men and the ways of men.’” Id. at 113 (quoting *Aetna Life Ins. Co. v. Ward*, 140 U.S. 76, 88 (1891)) (alteration in original).
- 43 See *Westcott v. Crinklaw*, 68 F.3d 1073, 1077 (8th Cir. 1995) (“expert testimony going to the issue of credibility is not admissible”); *United States v. Shay*, 57 F.3d 126, 131 (1st Cir. 1995) (“[A]n expert's opinion that another witness is lying or telling the truth is ordinarily inadmissible ... because the opinion exceeds the scope of the expert's specialized knowledge and therefore merely informs the jury that it should reach a particular conclusion.”).
- 44 Christian A. Meissner & Saul M. Kassin, “He's guilty!”: Investigator Bias in Judgments of Truth and Deception, 26 Law & Hum. Behav. 469, 470 (2002); see also Saul M. Kassin, *Human Judges of Truth, Deception, and Credibility: Confident But Erroneous*, 23 *Cardozo L. Rev.* 809, 811 (2002):
Surprisingly, however, professionals who regularly make these kinds of [truth-determination] judgments for a living, like the rest of us, are highly prone to error. In one study, researchers Paul Ekman and Maureen O'Sullivan were curious to know whether groups of so-called experts--such as police investigators; CIA, FBI, and military polygraph examiners; trial judges; psychiatrists; and U.S. Secret Service Agents--are truly better than the average person. Using stimulus materials from past studies--consisting of true and false stories--they found that college students had a 52.8 percent accuracy rate, which is pretty typical. Police detectives were only slightly higher, at 55.8 percent; CIA, FBI, and military polygraph examiners were at 55.7

JUNK SCIENCE AND THE EXECUTION OF AN..., 7 N.Y.U. J. L. &...

percent, trial judges were at 56.7 percent, and psychiatrists were at 57.6 percent. U.S. Secret Service Agent[s] won the prize, exhibiting a 64 percent accuracy rate, the highest of all groups.

45 Willingham Transcript, supra note 16, vol. XII, at 54.

46 Willingham Transcript, supra note 16, vol. XI, at 262 (“In my opinion, they are self-inflicted.”). Vasquez also testified that Willingham did not suffer smoke inhalation. Id. at 265. He had no firsthand knowledge of Willingham's condition immediately after the fire. He started his investigation on December 27, four days after the fire. Id. at 229.

47 See *Willingham v. State*, 897 S.W.2d 351, 358 (Tex. Crim. App. 1995) (“Johnny Webb, a State's witness, testified that appellant confessed to him that he committed the offense; that appellant explained in detail how he poured lighter fluid throughout the house, purposely burned one of the children, set the house on fire, fled, and refused to go back into the house to rescue the children.”).

48 Stephen S. Trott, *Words of Warning for Prosecutors Using Criminals as Witnesses*, 47 *Hastings L.J.* 1381, 1394 (1996); see also Robert M. Bloom, *Rattling: The Use and Abuse of Informants in the American Justice System* 65 (2002) (“[Leslie Vernon White] admitted to consistently fabricating confessions of fellow inmates and offering perjured testimony to courts.”); John Grisham, *The Innocent Man: Murder and Injustice in a Small Town 196-98* (2006) (discussing the snitch testimony of Terri Holland in the Ron Williamson case; Williamson was later exonerated by DNA); Northwestern Univ. Sch. of Law, Center on **Wrongful Convictions**, *The Snitch System* 3 (2004), [http:// www.innocenceproject.org/docs/SnitchSystemBooklet.pdf](http://www.innocenceproject.org/docs/SnitchSystemBooklet.pdf) (noting that snitch cases account for 45.9% of the 111 death row exonerations since the death penalty was restored in the 1970s; most were jailhouse informants); Vesna Jaksic, *Calif. May Crack Down on Use of Jailhouse Informants*, *Nat'l L.J.*, Dec. 20, 2006, http://www.law.com/jsp/nlj/PubArticleNLJ.jsp?id=900005469907&Calif_may_crack_down_on_use_of_jailhouse_informants (reporting that the California Commission on the Fair Administration of Justice issued guidelines on the use of jailhouse informants; and that of the 117 death penalty appeals pending in the California State Public Defender office, seventeen involved testimony by in-custody informants and six involved testimony by informants in constructive custody).

49 Facts on Post-Conviction DNA Exonerations, The Innocence Project, http://www.innocenceproject.org/Content/Facts_on_PostConviction_DNA_Exonerations.php, (last visited Feb. 6, 2012).

50 Willingham Transcript, supra note 16, vol. XI, at 13-14, 26-27 (testimony of Johnny Webb).

51 Id. at 23.

52 As one court wrote: “We are not unaware of the reality that the Government has ways of indicating to witness's counsel the likely benefits from cooperation without making bald promises” *United States v. Ramirez*, 608 F.2d 1261, 1266 n.9 (9th Cir. 1979) (citing *United States v. Butler*, 567 F.2d 885, 888 (9th Cir. 1978)); see also R. Michael Cassidy, “Soft Words of Hope:” *Giglio, Accomplice Witnesses, and the Problem of Implied Inducements*, 98 *Nw. U. L. Rev.* 1129, 1132 (2004) (“The Court's decision in *Giglio* has created an incentive for prosecutors to make representations to an accomplice witness that are vague and open-ended, so that they will not be considered a firm ‘promise’ mandating disclosure.... Such indefinite agreements have the added advantage of allowing prosecutors to argue to the jury that no specific promise has been made to the witness; this is viewed as tactically more advantageous to the government because it prevents the factfinder from second-guessing the appropriateness of concessions ultimately conferred.”). Another authority put it this way:

To enhance the credibility of his testimony, an informant often testified that there have been no promises of benefits made to them in return for their testimony. Even though nothing may be explicitly stated, both the prosecutor and the informant knew that there will be some compensation for the testimony. “The practice (of promising rewards) was done by a wink and a nod and it was never necessary to have any kind of formal understanding.”

Bloom, supra note 48, at 66 (citing Los Angeles County Grand Jury, *Investigation of the Involvement of Jail House Informants in the Criminal Justice System in Los Angeles County* 39 (1990); Ted Rohrlich, *Perjurer Sentenced to 3 Years; Crime: Informant Blew the Whistle on Use of Jailhouse Liar-for-Hire, but No Law Officers Were Charged for Conspiring with Him*, *L.A. Times*, May 20, 1992, at B1 (quoting Douglas Dalton, special counsel to the Los Angeles County Grand Jury)).

JUNK SCIENCE AND THE EXECUTION OF AN..., 7 N.Y.U. J. L. &...

- 53 Grann, supra note 1, at 52.
- 54 Willingham Transcript, supra note 16, vol. XI, at 18 (testimony of Johnny Webb).
- 55 Grann, supra note 1, at 52.
- 56 See [Willingham v. State](#), 897 S.W.2d 351, 354-55 (Tex. Crim. App. 1995):
Neighbors of appellant testified that as the house began smoldering, appellant was “crouched down” in the front yard, and despite the neighbors' pleas, refused to go into the house in any attempt to rescue the children.... The testimony at trial demonstrates that appellant neither showed remorse for his actions nor grieved the loss of his three children. Appellant's neighbors testified that when the fire “blew out” the windows, appellant “hollered about his car” and ran to move it away from the fire to avoid its being damaged. A fire fighter also testified that appellant was upset [the next day] that his dart board was burned. One of appellant's neighbors testified that the morning following the house fire, Christmas Eve, appellant and his wife were at the burned house going through the debris while playing music and laughing.
In prison, Willingham said he moved the car because he was afraid it would catch on fire and explode. Grann, supra note 1, at 50. See also Mills & Possley, supra note 4 (one juror “said she would have found Willingham guilty even without the **arson** finding solely because he did not try to save his children.”).
- 57 Willingham transcript, supra note 16, vol. XIII, at 43-44.
- 58 Willingham Transcript, supra note 16, vol. XI, at 72 (testimony of Dianne Barbe, neighbor of Willingham). See also id. at 88 (testimony of Brandy Barbe, neighbor of Willingham) (“He was screaming that there was fire, that his babies were burning and for someone to help him, to call 911.”); id. at 103 (testimony of Buffy Barbe, neighbor of Willingham) (“He was hollering, ‘My babies are inside burning up. Help me.’”).
- 59 Id. at 128, 132 (testimony of Ronald Franks).
- 60 Id. at 149 (testimony of Jason Grant).
- 61 Id. at 104 (testimony of Buffy Barbe, neighbor of Willingham).
- 62 The fire “experts who reviewed the case didn't put any stock in the claims that Willingham's behavior was damning. They say experience shows that there is no way to predict how people will react in a fire or to the grief of losing loved ones.” Mills & Possley, supra note 4. The literature on rape trauma syndrome also illustrates this point. There is no typical way that a rape victim will react. Some victims are hysterical; others are calm. See Giannelli & Imwinkelried, supra note 13, § 9.03 (discussing rape trauma syndrome).
- 63 “The witnesses' testimony also grew more damning after authorities had concluded, in the beginning of January, 1992, that Willingham was likely guilty of murder. In Diane Barbee's initial statement to authorities, she had portrayed Willingham as ‘hysterical,’ and described the front of the house exploding. But on January 4th, after **arson** investigators began suspecting Willingham of murder, Barbee suggested that he could have gone back inside to rescue his children, for at the outset she had seen only ‘smoke coming from out of the front of the house’--smoke that was not ‘real thick.’” Id. at 49-50.
- 64 This type of contextual bias is not limited to witnesses; everyone is subject to it, even professionals. See Itiel E. Dror et al., Contextual Information Renders Experts Vulnerable to Making Erroneous Identifications, 156 Forensic Sci. Int'l 74 (2006); D. Michael Risinger et al., [The Daubert/Kumho Implications of Observer Effects in Forensic Science: Hidden Problems of Expectation and Suggestion](#), 90 Calif. L. Rev. 1, 38-39 (2002).
- 65 In the Willis case, discussed below, the “[w]itnesses maintained that Willis had acted suspiciously; he moved his car out of the yard, and didn't show ‘any emotion,’ as one volunteer firefighter put it.” Grann, supra note 1, at 56. Similarly, in the famous Lime Street fire, witnesses “told authorities that Lewis seemed too calm during the fire and had never tried to get help.” Id. at 59. In both cases, the defendants were exonerated.

JUNK SCIENCE AND THE EXECUTION OF AN..., 7 N.Y.U. J. L. &...

- 66 [Id. at 47.](#)
- 67 Willingham Transcript, supra note 16, vol. XIV, at 21 (testimony of Stacy Willingham).
- 68 Willingham Transcript, supra note 16, vol. XI, at 18 (testimony of Johnny Webb).
- 69 After the trial, the prosecution would suggest another motive: “[A]s the local district attorney, Pat Batchelor, put it, ‘The children were interfering with his beer drinking and dart throwing.’” Grann, supra note 1, at 47. This appears to be no more than speculation.
- 70 The prosecution called her as a hostile witness in the penalty stage. She testified that Todd “never hurt those kids.” Willingham Transcript, supra note 16, vol. XIV, at 5. Then, the prosecutor asked: “Well, are you the one who hurt the kids?” *Id.* There is no **evidence** in the record that anyone had ever hurt the children. The prosecutor also cross-examined her regarding life insurance. *Id.* at 20-21.
- 71 “Dunn [defense counsel] told me that Willingham had wanted to testify, but Martin and Dunn thought that he would make a bad witness.” Grann, supra note 1, at 48.
- 72 At most, it made Willingham appear callous and perhaps a coward, if one believes that a father should have entered the burning house. Willingham told investigators that the smoke was too thick and that he was singed by flames. See also Grann, supra note 1, at 63 (Later Willingham “confessed to his parents that there was one thing about the day of the fire he had lied about. He said that he had never actually crawled in the children’s room. ‘I just didn’t want people to think I was a coward,’ he said. [Dr.] Hurst told me, ‘People who have never been in a fire don’t understand why those who survive often can’t rescue the victims. They have no concept of what a fire is like.’”).
- 73 The penalty phase included its own version of junk science. In this phase, the prosecution offered the testimony of Dr. James Grigson, who was known as “Dr. Death.” Ron Rosenbaum, *Travels With Dr. Death and Other Unusual Investigations* 206 (1991). Grigson testified that Willingham was a violent sociopath without ever interviewing him. One scholar labeled Grigson’s testimony in death penalty cases as “at the brink of quackery.” George E. Dix, *The Death Penalty, “Dangerousness,” Psychiatric Testimony, and Professional Ethics*, 5 *Am. J. Crim. L.* 151, 172 (1977). As a prominent conservative author noted: One could favor the death penalty and “yet still recoil at the thought that a junk science fringe of psychiatry ... could decide who will be sent to the gallows.” Peter W. Huber, *Galileo’s Revenge: Junk Science in the Courtroom* 220 (1991).
- 74 [Willingham, 897 S.W.2d at 354.](#) The United States Supreme Court denied certiorari from the direct appeal. [Willingham v. Texas, cert. denied, 516 U.S. 946 \(1995\)](#) (direct appeal).
- 75 The Texas Court of Criminal Appeals denied post-conviction relief in 1997. *Ex parte Willingham*, No. 35, 162 (Tex. Crim. App. 1997). Federal habeas challenges were also unsuccessful. [Willingham v. Cockrell, No. 02-10133, 2003 WL 1107011 \(5th Cir. Feb. 17, 2001\)](#), cert. denied, *Willingham v. Dretke*, 524 U.S. 917 (1998).
- 76 Report of Dr. Gerald Hurst, *Ex parte Cameron Todd Willingham*, Trial Court No. 24, 4670, (Tex Dist. Ct. Feb 13, 2004), available at <http://www.pbs.org/wgbh/pages/frontline/death-by-fire/documents/hursts-2004-report.html>.
- 77 Willingham’s lawyer also petitioned the Court of Criminal Appeals, which ruled that the Hurst report was not newly discovered **evidence**.
- 78 Grann, supra note 1, at 62 (“The Innocence Project obtained, through the Freedom of Information Act, all the records from the governor’s office and the board pertaining to Hurst’s report. ‘The documents show that they received the report, but neither office has any record of anyone acknowledging it, taking note of its significance, responding to it, or calling any attention to it within the government,’ Barry Scheck said.”). See also Dave Mann, *Fire and Innocence*, *Tex. Observer*, (Dec. 3, 2009), www.texasobserver.org/fire-and-innocence/ (“Because the governor’s office has refused to release relevant documents, it’s unclear what, if anything, the governor’s staff did with Hurst’s report or whether Perry ever saw it.”).

JUNK SCIENCE AND THE EXECUTION OF AN..., 7 N.Y.U. J. L. &...

- 79 “[A] certification that a government entity exists and an appropriate process is in place to conduct independent external investigations into allegations of serious negligence or misconduct substantially affecting the integrity of the forensic results committed by employees or contractors of any forensic laboratory system, medical examiner's office, coroner's office, law enforcement storage facility, or medical facility in the State that will receive a portion of the grant amount.” 42 U.S.C. §3797k(4) (2004).
- 80 See Giannelli, supra note 11.
- 81 See Irma Rios Et. Al., Quality Assurance Audit of Houston Police Dep't Crime Laboratory-DNA/Serology Section (2002) (revealing a dysfunctional organization with serious contamination issues and an untrained staff using shoddy science) (on file with the New York University Journal of Law and Liberty). See also Nick Madigan, Houston's Troubled DNA Crime Lab Faces Growing Scrutiny, N.Y. Times, Feb. 9, 2003, at A20 (reporting that operations were suspended after an audit found numerous problems, including poor calibration and maintenance of equipment, improper record keeping, and a lack of safeguards against contamination; “Among other problems, a leak in the roof was found to be a potential contaminant of samples on tables below.”); Giannelli, supra note 11, at 187-91 (discussing the Houston crime lab scandal).
- 82 Rodney Ellis, Editorial, Want Tough on Crime? Start by Fixing HPD Lab., Houston Chron., Sept. 5, 2004. Similarly, the chair of the legislative committee investigating the lab stated: “It's a comedy of errors, except it's not funny.” Adam Liptak, Houston DNA Review Clears Convicted Rapist, and Ripples in Texas Could be Vast, N.Y. Times, March 11, 2003 (quoting state Representative Kevin Bailey).
- 83 Third Report of the Independent Investigator for the Houston Police Dep't Crime Laboratory and Property Room 5 (June 30, 2005).
- 84 Tex. Crim. Proc. Code art. 38.01(4)(a)(3) (2005) (among other duties, the Commission should “investigate, in a timely manner, any allegation of professional negligence or misconduct that would substantially affect the integrity of the results of a forensic analysis conducted by an accredited laboratory, facility, or entity.”).
- 85 See Mills & Possley, supra note 4.
- 86 Two of the experts are authors of standard texts in the field. See John DeHann & David J. Icove, Kirk's Fire Investigations (7th ed. 2011); Lentini, supra note 17. The third expert, Kendall Ryland, was the Louisiana fire chief. Phillip Martin, Juror in Willingham Execution Case: “Maybe This Man Was Innocent”, Burnt Orange Report, Oct. 7, 2009.
- 87 Carpenter et al., supra note 30, at 3. One of these experts, John Lentini, was also one of the experts consulted by the Chicago Tribune. Maurice Possley, Report: Inmate Wrongly Executed; Arson Experts Say Evidence in Texas Case Scientifically Invalid, Chi. Trib., May, 3, 2006, at C1.
- 88 Willingham Transcript, supra note 16, vol. XI, at 228 (“With the exception of a few, most all of them.”).
- 89 Carpenter et al., supra note 30, at 5-6.
- 90 Willingham Transcript, supra note 16, vol. XI, at 228 (“Unfortunately, fires injure a lot of people, kill a lot of people. It's about 50 percent.”).
- 91 Carpenter et al., supra note 30, at 5-6.
- 92 See Mary Alice Robbins, New-York Based Innocence Project Attacks Texas Arson Convictions, 22 Tex. Lawyer, May 8, 2006.
- 93 See Willis v. Cockrell, No. P-01-CA-20, 2004 WL 1812698, at *1 (W.D. Tex. Aug. 09, 2004) (finding, among other deficiencies, ineffective assistance of counsel).
- 94 “Ori T. White, then the district attorney in Fort Stockton, filed a certificate of actual innocence for Willis in the Court of Criminal Appeals.” Robbins, supra note 92.

JUNK SCIENCE AND THE EXECUTION OF AN..., 7 N.Y.U. J. L. &...

- 95 TFSC Report, *supra* note 39, at 13.
- 96 See Letter from Innocence Project, to Texas Forensic Science Comm'n (Aug. 20, 2010) (on file with the New York University Journal of Law and Liberty).
- 97 See Mary Alice Robbins, Lack of Money, Members Stalled Launch of Crime Lab Commission, 22 Tex. Lawyer, Oct. 23, 2006 (“It’s obvious that somebody’s dragging their feet on this,” says state Sen. Juan Hinojosa, Senate sponsor of H.B. 1068.”). Under the statute, the Governor appoints four members--two forensic science experts, a prosecutor, and a defense attorney. The Lieutenant Governor appoints three members--one each from University of Texas and Texas A&M University, specializing in clinical laboratory medicine, and one from Texas Southern University, specializing in pharmaceutical laboratory research. See [Tex. Crim. Proc. Code art. 38.01\(3\)\(a\)](#) (2005).
- 98 The Commission voted to investigate the Willingham case at its August 15, 2008 meeting.
- 99 Craig L. Beyler, Analysis of the Fire Investigation Methods and Procedures Used in the Criminal [Arson](#) Cases Against Ernest Ray Willis and Cameron Todd Willingham, Aug. 17, 2009, at 51. Still another expert agreed. See Letter from Mark Goodson, Goodson Engineering, to Texas Forensic Science Comm'n (Sept. 23, 2009), at 1 (“I first off want to commend Dr. Beyler for an excellent report. His findings, in my view, are accurate. Moreover, the report agrees with the previous reports by both Lentini, Carpenter et al. and Dr. Hurst.”).
- 100 See Christy Hoppe, Perry Defends Removing 3: He Says He's Following Protocol, but Critics Believe He's Derailing [Arson](#) Inquiry, *Dall. Morning News*, Oct. 2, 2009, at 3A; Robbins, *Fired Up*, *supra* note 7 (“[Former Commissioner] Levy says he believes things went south for the commission after [former Chair] Bassett released Beyler's report to the public in August as he was required by law to do.”). The meeting was scheduled for October 2, 2009. Grann's article had been published several weeks earlier. See Grann, *supra* note 1.
- 101 Hoppe, *Perry Defends Removing 3*, *supra* note 100 (noting that Bradley was “known as one of the toughest law-and order prosecutors in the state”).
- 102 See Emily, *supra* note 4 (“Perry's replacements were seen by some as a political maneuver intended to change the outcome of the commission's decision.”); Christy Hoppe, Perry Ousts Officials Before [Arson](#) Hearing: He's Assailed as New Chair Delays Session on Flawed Case that Led to Execution, *Dall. Morning News*, Oct. 1, 2009, at 1A; David Mann, Fire and Innocence, *Tex. Observer*, Dec. 2, 2009 (“Then in late September, Perry booted three members off of the Texas Forensic Science Commission, which was investigating the Willingham and Willis cases, just three days before a crucial hearing on scientists' findings. Perry's new appointees promptly canceled the hearing and have yet to reschedule it. Even conservative commentators cried cover-up, suggesting that Perry, in a tough battle for re-election, was trying to subvert an investigation that might prove he oversaw the execution of an innocent man.”).
- 103 See Robert T. Garrett, Forensics Panel Faulted: Former Chairman Says Overhaul May Delay [Arson](#) Review for Years, *Dall. Morning News*, Nov. 12, 2009 (Bradley faulted former chairman Sam Bassett, saying he “utterly failed to adopt even a definition of negligence or misconduct.”); CSI: Texas: Governor Shakes Up Commission, Covers Tracks, *Hous. Chronicle*, Nov. 17, 2009 (Bradley “canceled the meeting [with fire expert Craig Beyler] and raised a number of issues about the commission's lack of rules and procedures.”).
- 104 See Emily, *supra* note 4 (“The Texas Forensic Science Commission rebelled Friday against its head commissioner, refusing to accept his draft report clearing [arson](#) investigators of misconduct or negligence in a 1991 fatal fire where flawed science was used to determine the blaze was intentionally set.”).
- 105 See Rick Casey, The Revolt of the Scientists: No Legalistic Briar Patch Allowed, *Hous. Chronicle*, Jan. 31, 2010, at B1. See also Rick Casey, A Win for Bradley, and Another Loss: Panel Worried About Funding, *Hous. Chronicle*, Feb. 3, at B1 (“[T]he seven scientists on the nine-member commission rebelled at a set of policies and procedures presented by Bradley that would have given him formidable powers as chairman, including naming members and chairs of three standing committees and of ad-hoc committees that will direct the investigations of alleged failures at police labs and other agencies.”).

JUNK SCIENCE AND THE EXECUTION OF AN..., 7 N.Y.U. J. L. &...

- 106 Christy Hoppe & Gromer Jeffers Jr., [Arson](#) Expert: Furor over Texas Forensics Masks Dangerous Trend, *Dall. Morning News*, Oct. 15, 2009. See also Steve Mills, *Texas Execution: Statements by Gov. Rick Perry Others Don't Align with Facts*, *Chi. Trib.*, Oct. 19, 2009, at C12.
- 107 Letter from Terry Jacobson, City Attorney, to Leigh Tomlin, Comm'n Coordinator, *Tex. Forensic Sci. Comm'n* (Aug. 20, 2010) (“[T]he commission lacks the authority to even review this matter.”) [hereinafter SFMO Letter].
- 108 Rick Casey, *Willingham: Scientists vs. Lawyers*, *Hous. Chronicle*, Jan. 10, 2010.
- 109 Erin Mulvaney, *National Experts Criticize States's Study of Fatal 1991 House Fire*, *Dall. Morning News*, Jan. 8, 2011. See also Aziza Musa, [Arson](#) Experts Testify in Willingham Investigation, *Tex. Trib.*, Jan. 7, 2011 (“Beyler accused the original investigators of ignoring eyewitnesses, whose testimony contradicted the [arson](#) determination.”); Allan Turner, [Arson](#) Probe that Led to Corsicana Man's Execution Assailed, *Hous. Chronicle*, Jan. 8, 2011 (“Beyler also faulted Vasquez for failing to investigate the possibility that the fire had been set by one of the children or by an intruder. Rather than systematically explore possible causes such as an electrical short, Coricana authorities ‘shoveled out the room and put it out the window,’ Beyler charged.”).
- 110 See supra note 86.
- 111 Mulvaney, supra note 109.
- 112 Casey, supra note 108.
- 113 By this time, Willingham's family along with the Innocence Project sought a judicial inquiry into Willingham's conviction. *Petition to Convene a Court of Inquiry and For Declaration to Remedy Injury to Mr. Willingham's Reputation under the Texas Constitution*, 299th Dist. Court, Travis County, Tex., Sept. 24, 2010. See also [Tex. Crim. Proc. Code art. 52.01](#) (1995) (courts of inquiry). The prosecutors responded by asking the judge to recuse himself and then filed a writ of mandamus on the recusal issue. *Relator's Petition for Writ of Mandamus and Emergency Motion for Immediate Stay*, 3d Jud. Dist. Court of Appeals, Oct. 14, 2010.
- 114 Letter from John Bradley, *Tex. Forensic Sci. Comm'n*, to Greg Abbott, Attorney Gen., *Tex.* (Aug. 28, 2010) (on file).
- 115 Controversially, the Attorney General eventually opined that the Commission lacked jurisdiction. See [Tex. Att'y Gen. Op. No. GA-0866](#), at *4 (2011), [2011 WL 4352017](#) (concluding that the TFSC cannot investigate incidents before September 1, 2005, long after the Willingham incident). The TFSC would eventually accept the Attorney General's reading of the statute. See *Tex. Forensic Sci. Comm'n, Addendum to the April 15, 2011 Report of the Texas Forensic Science Commission, Willingham/Willis Investigation*, at 3-4 (Oct. 28, 2011), available at <http://content.news8austin.com/auscontent/files/WillinghamWillisReport.pdf>. See also Ryan M. Goldstein, Note, [Improving Forensic Science Through State Oversight](#), *90 Tex. L. Rev.* 225, 245 (2011) (referring to the AG opinion as a “narrow interpretation” of the jurisdictional statute).
- 116 TFSC Report, supra note 39, at 39. Other recommendations included enhanced certification, comprehensive reports, retention of records, and standards in testifying. *Id.* at 39-52.
- 117 *Id.* at 21-28.
- 118 *Id.* at 36.
- 119 *Id.* at 48.
- 120 Letter from Paul Maldonado, State Fire Marshal, to Leigh Tomlin, Comm'n Coordinator, *Tex. Forensic Sci. Comm'n* (Aug. 20, 2010) (on file with the *New York University Journal of Law and Liberty*) (emphasis added).
- 121 TFSC Report, supra note 39, at 16.

JUNK SCIENCE AND THE EXECUTION OF AN..., 7 N.Y.U. J. L. &...

- 122 Id. at 41.
- 123 “Even Edward Cheever, one of the state deputy fire marshals who had assisted in the original investigation of the 1991 fire, acknowledged that Hurst's criticism was valid. ‘At the time of the Corsicana fire, we were still testifying to things that aren't accurate today,’ he said. ‘They were true then, but they aren't now. ‘Hurst,’ he added, ‘was pretty much right on.... We know now not to make those same assumptions.’” Mills & Possley, *supra* note 4. After the issue became politicized, Cheever “told the [Texas Forensic Science Commission] he had been misquoted by a Chicago Tribune reporter who wrote that he admitted the standards Vasquez used now may be considered outdated.” Turner, *supra* note 109. Assistant Chief Fogg, however, has not changed his opinion. See *Frontline: Death by Fire*, *supra* note 5.
- 124 Texas officials have suggested that Willingham was guilty, even if the **arson evidence** was flawed. The statements of Willingham's wife, Stacy, are cited as support for this view. At the trial, she testified at the penalty stage that she believed him to be innocent. See *supra* note 70. After the trial, she worked for his exoneration. See Grann, *supra* note 1, at 47, 49 (Stacy “wrote to Ann Richards, then the governor of Texas, saying, ‘I know him in ways that no one else does when it comes to our children. Therefore, I believe that there is no way he could have possibly committed this crime.’” She reported to “investigators that even though Willingham hit her he had never abused the children--‘Our kids were spoiled rotten,’ she said--and she did not believe that Willingham could have killed them.”). She later remarried. For over a decade, she did not visit him on death row. At some point she changed her mind. Recently, she said that he had confessed to her in their last meeting immediately before his execution. Yet, “[i]n their final meeting ... he did not confess, she told the Tribune.” Mills & Possley, *supra* note 4.
- 125 See SFMO Letter, *supra* note 107 (“The guidelines NFPA 921 set out were used by the State Fire Marshal prior to NFPA 921's initial publication”; “The SFMO staff began referencing and received training on NFPA 921 almost immediately after its initial publication in 1992.”).
- 126 *Herrera v. Collins*, 506 U.S. 390, 415 (1993). See also *Kansas v. Marsh*, 548 U.S. 163, 193 (2006) (Scalia, J., concurring) (“Reversal of an erroneous conviction on appeal or on habeas, or the pardoning of an innocent condemnee through executive clemency, demonstrates not the failure of the system but its success. Those devices are part and parcel of the multiple assurances that are applied before a death sentence is carried out.”); *Dretke v. Haley*, 541 U.S. 386, 399 (2004) (Kennedy, J., dissenting) (“Among its benign if too-often ignored objects, the clemency power can correct injustices that the ordinary criminal process seems unable or unwilling to consider”).
- 127 Grann, *supra* note 1, at 62 (“The vote was unanimous.... [T]he board deliberates in secret, and its member are not bound by any specific criteria. The board members did not even have to review Willingham's materials, and usually don't debate a case in person; rather, they cast their votes by fax--a process that has become known as ‘death by fax.’”).
- 128 *Tex. Gov't Code Ann. § 508.047(b)* (West 2004) (“The members of the board are not required to meet as a body to perform the members' duties in clemency matters”). In a 2002 case, a criminal defendant alleged that “only one live clemency hearing has been held in the past thirty years.” *Lagrone v. Cockrell*, No. Civ.A.4:99-CV-0521-G, 2002 WL 1968246, at *23 (N.D. Tex. Aug. 19, 2002).
- 129 Grann, *supra* note 1, at 62. Texas's lackluster reputation for reviewing death-penalty cases includes the conduct of Sharon Keller, presiding judge of the Court of Criminal Appeals, who refused to keep the clerk's office open past five o'clock to permit a last-minute petition by an inmate executed later that night. See Gretel C. Kovach, *Mixed Opinions of a Judge Accused of Misconduct*, N.Y. Times, Mar. 7, 2009, at A14.
- 130 *Ex parte Tucker*, 973 S.W.2d 950, 951 (Tex. Crim. App. 1998) (Overstreet, J., concurring). See also Steve Woods, Comment, *A System Under Siege: Clemency and the Texas Death Penalty After the Execution of Gary Graham*, 32 *Tex. Tech L. Rev.* 1145, 1146 (2001) (“Critics are enraged at the Texas Board of Pardons and Paroles's record of only recommending one individual for clemency since 1995.... The execution of Gary Graham in Huntsville, Texas, in June of 2000, cast a worldwide spotlight on Texas and its clemency procedure and has cemented the need for the state to examine its methods to determine whether any improvements can, or even should, be made.”).

JUNK SCIENCE AND THE EXECUTION OF AN..., 7 N.Y.U. J. L. &...

- 131 Until recently, only New York had established a Commission on Forensic Science. It is authorized to (1) develop minimum standards and a program of accreditation for all state laboratories, (2) establish minimum qualifications for laboratory directors and other personnel, and (3) approve forensic laboratories for the performance of specific forensic methodologies. [N.Y. Exec. Law § 995-b \(McKinney 2012\)](#). See Giannelli, *supra* note 11, at 170.
- 132 See generally Mann, *supra* note 78 (discussing the cases of Curtis Severns, Ed Graf, and Alfredo Guardiola).

7 NYUJLL 221

End of Document

© 2017 Thomson Reuters. No claim to original U.S. Government Works.

4 Va. J. Crim. L. 1

Virginia Journal of Criminal Law
Spring, 2016

Article

M. Chris Fabricant¹ Tucker Carrington²

Copyright © 2016 by Virginia Criminal Law Association; M. Chris Fabricant, Tucker Carrington

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE EVOLUTION FROM MAGIC TO LAW

ABSTRACT

*A decade ago, a controversial article in Science Magazine predicted a coming “paradigm shift” that would push forensic sciences toward fundamental change as the result of “[l]egal and scientific forces ... converging to drive an emerging skepticism about the claims of the traditional forensic individualization sciences.”³ This article argues that the predicted paradigm shift has occurred. We support our thesis through a deconstruction of the jurisprudence of two of the forensic disciplines implicated in numerous **wrongful convictions**--forensic odontology (bite mark analysis) and forensic hair microscopy--and an examination of a confluence of unprecedented events currently altering the landscape of forensic sciences. The empirical **evidence** and data gathered here demonstrates that traditional forensic identification techniques, as well as *2 the doctrines supporting them, are ultimately no more than a house of cards built on unvalidated hypotheses and unsubstantiated or non-existent data. Several far-reaching consequences result. Among those consequences are that state--and, to some extent, federal--jurisprudence that stands for the proposition that this type of **evidence** is admissible is objectively erroneous and must be reevaluated and effectively rejected as invalid precedent. The long-overdue paradigm shift presents a unique ethical challenge to criminal justice professionals, one that current professional ethics regimes fail to adequately capture, even though fundamental due process norms compel the conclusion that prosecutors, defense attorneys, forensic experts, and their respective governing bodies have an ethical, moral, and legal obligation to revisit affected cases and provide appropriate remedies. Put differently, the “path forward” for forensic sciences that the National Academy of Sciences identified in its seminal 2009 report must have a rear-view mirror.⁴*

TABLE OF CONTENTS

PREFACE	5
INTRODUCTION	12
I. “THE SHIFTED PARADIGM”	15
A. FORENSIC EVIDENCE-BASED POST-CONVICTION EXONERATIONS	19
B. NATIONAL ACADEMY OF SCIENCES REPORT AND RECENT COMPLEMENTARY FORENSIC DEVELOPMENTS	23
1. National Commission on Forensic Science	24
2. Basic and Applied Research and Standards Development	26

C. FEDERAL CASE AUDIT & STATE LEGISLATION	28
D. EMPIRICAL DATA	36
II. BITE MARK EVIDENCE	38
A. INTRODUCTION	38
B. THE RISE OF BITE MARK MATCHING	42
III. HAIR EVIDENCE: UNCHALLENGED AND UNVALIDATED	63
A. INTRODUCTION	63
B. THE RISE OF HAIR MICROSCOPY	71
C. THE FBI'S TRAINING OF HUNDREDS OF STATE HAIR EXAMINERS TO PROVIDE FALSE AND MISLEADING TESTIMONY	87
D. THE FALL OF HAIR MICROSCOPY	91
IV. ATTENDANT OBLIGATIONS	93
A. UNIQUE NATURE OF ETHICAL & PROFESSIONAL OBLIGATIONS	99
B. SUGGESTED SOLUTIONS	109
C. MODIFICATION OF PROCEDURAL BARRIERS	110
CONCLUSION	112

*5 PREFACE

When a federal magistrate judge recommended that the United States District Court for the Middle District of Pennsylvania exonerate Han Tak Lee for the murder by arson of his young daughter, he began his report this way: “Slow and painful has been man's progress from magic to law.”⁵ Lee's daughter, the court explained, had perished “in a tragic cabin fire at a religious retreat,” and the State's evidence:

was based, to a substantial degree, upon what was at the time undisputed scientific evidence concerning the source and origin of this fire, fire origin evidence which tended to show that the fire which consumed this cabin and took the life of ... [the victim] was deliberately set by the defendant in a calculated fashion.⁶

Lee had been wrongly imprisoned for twenty-five years, and the State's conviction rested on the theory--elicited through expert testimony--that Lee “was especially cruel and calculating, dousing ... [the] small cabin in Pennsylvania's Pocono Mountains with more than 60 gallons of gasoline and heating fuel and setting at least eight fires, ending at the front door to block any chance of escape.”⁷

When the court recommended that Lee be freed, however, it left no room for debate about either Lee's innocence or the character of the evidence that had claimed a quarter century of his life. “Today,” the court wrote:

*6 with the benefit of extraordinary progress in human knowledge regarding fire science over the past two decades it is now uncontested that this fire science evidence - which was a critical component in the quantum of proof that led to ... [the] conviction - is invalid, and that much of what was presented to ... [the] jury as science is now conceded to be little more than superstition.⁸

Han Tak Lee is an important case. This article argues that its primary importance rests not, however, in its innocence narrative, which is now, unfortunately, a familiar one.⁹ Nor does it rest even in the sub-narrative of its particular taxonomy: the set of wrongful convictions whose root cause is flawed forensic science.¹⁰ Instead, its significance derives from the court's searing and ultimately dispositive critique of the forensic discipline that formed the core of the wrongful

conviction--arson science--as well as of the false accord that has been granted to certain forensic disciplines in our criminal justice system over time. Despite their widespread acceptance by criminal courts, such disciplines are, in final analysis, mere “superstition.”¹¹

The judicial critique--which elevated scientific data above doctrinal dogma-- was a long time in coming and necessarily calls into question the force and legitimacy of precedent as basis to introduce purportedly scientific **evidence**. It has been a decade since an article in *Science Magazine* predicted what the authors termed the coming *7 “paradigm shift”¹²--forensic science's evolution away from magic and toward law--the so-called “shot heard round the forensic science community.”¹³ Based on a historical assessment of the jurisprudence concerning two forensic assays--bite mark identification and hair microscopy--as well as their underlying empirical **evidence** and data, this article presents a new, parallel narrative: the failure of courts and litigants to distinguish between magic and science in the first instance, and the judicial system's continuing reflexive reliance on deeply flawed, scientifically invalid precedent to support the admissibility of false and misleading **evidence**.

Consider, for example, the relationship between the 1985 Wisconsin conviction of Robert Lee Stinson for the murder of his elderly neighbor and the 1992 Mississippi conviction of Levon Brooks for the sexual assault and murder of a three-year old girl. The only direct **evidence** against Stinson was the bite mark testimony of two board-certified “Diplomates” of the American Board of Forensic Odontology (ABFO). One expert concluded that bite marks on the victim “had to have been made by teeth identical”¹⁴ to Stinson's, and that there was “no margin for error”¹⁵ in his conclusion. The other expert concurred, testifying the bite mark **evidence** was “high quality”¹⁶ and “overwhelming.”¹⁷ In Brooks' case, the State also presented the testimony of a board-certified forensic odontologist, who, utilizing a purportedly path-breaking new forensic technique, testified that the only direct *8 **evidence** linking Brooks to the crime--a series of bite marks on the victim--matched” Brooks' teeth in such a way that “it could be no one but Levon Brooks that bit this girl's arm.”¹⁸ Stinson and Brooks were both exonerated after DNA testing proved that they were not the perpetrators--Stinson in 2009¹⁹ and Brooks in 2008.²⁰

Apart from sharing the same type of inculcating forensic **evidence**, *Stinson* and *Brooks* would appear to have little in common: the convictions are separated by nearly a decade; the crimes occurred on opposite sides of the country in jurisdictions that applied different standards to evaluate and admit expert testimony from two different experts.²¹ But a comparison of the underlying data tells a different story altogether: that the cases are actually strikingly similar, even codependent. In fact, each depended on the other for the failures of justice that occurred, both at trial and during the years-long, protracted struggle that Stinson and Brooks endured while seeking to prove their innocence.

In reviewing Brooks' conviction--and, more specifically, the propriety of admitting the bite mark **evidence**--the Mississippi Supreme *9 Court relied upon the reasoning in *Stinson* and other similar precedent,²² to not only find no error in the admission of the **evidence** and affirm the conviction, but also to issue a blanket pronouncement that “bite-mark identification is admissible in Mississippi” in an effort to preclude further challenges.²³ Nevertheless--in spite of both the exonerations of Stinson and Brooks and the overwhelming proof that the bite mark **evidence** presented in *Brooks* was spurious²⁴--*State v. Brooks and Stinson*, continue to stand as reliable authority for the wholesale admissibility of this branch of forensic science in Mississippi and Wisconsin state courts.

This kind of self-serving, court-facilitated pseudo-jurisprudence not only facilitates trial courts' wholesale admission of flawed **evidence**, but also insulates such decisions from appellate review, no matter how legally indefensible and intellectually dishonest. Post-conviction courts, moreover, typically avoid any rigorous analysis of a discipline's validity

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE..., 4 Va. J. Crim. L. 1

or of the propriety of a trial court's admissibility decision by invoking procedural bars.²⁵ In *Han Tak Lee*'s case, for example, post-conviction courts' review was primarily focused not on the analysis of fire science, but, instead, on procedural hurdles that Lee's request for post-conviction relief were required to overcome in state and federal *habeas corpus* litigation.²⁶ After the Pennsylvania Superior Court declined to address the *10 scientific legitimacy of the fire science,²⁷ Lee was denied review because his "claim of newly discovered evidence [that the fire evidence was not based on sound science] is not cognizable under § 2254 [state remedies in federal court]²⁸ because claims of actual innocence based on newly discovered evidence are never grounds for federal habeas relief absent an independent constitutional violation."²⁹

Han Tak Lee, like *Brooks* and *Stinson*, has been resolved.³⁰ But the scientifically invalid evidence and erroneous jurisprudence that led to those convictions continues to frustrate the truth-seeking mission of the criminal justice system, precluding legitimate claims of innocence in dozens-- perhaps hundreds--of other similarly situated cases, including capital convictions.³¹ The jurisprudence connecting *Han Tak Lee*, *Brooks*, *11 and *Stinson* to other, unresolved cases is not attenuated; it is direct and, as the evidence gathered here demonstrates, obvious. *Howard v. State*³² is a timely example. *Howard* is a pending death penalty case where the conviction rests almost exclusively on bite mark evidence.³³ The Mississippi Supreme Court--in affirming Howard's second conviction and refusing to grant him relief thus far³⁴--has relied on *Brooks* for substantive support of the discipline's validity, even though the case is a notorious incident of wrongful convictions and the same forensic expert who testified falsely in *Brooks* also testified in *Howard*.³⁵ Entirely absent from this appellate review is any discussion of the trial court's failure to conduct a rigorous analysis of the discipline before allowing it to be proffered to a capital jury as scientific evidence of guilt, as required by relevant case law and evidentiary rules.³⁶ Nor is there, in the court's reliance on *Brooks* for support, any recognition that those other courts' analysis was similarly deficient. Instead, the State of Mississippi continues to elevate procedural rules over scientific reality, arguing that Howard's claims are "barred from consideration both by the successive petition bar"³⁷ and "res judicata."³⁸ Thus far, the Mississippi Supreme Court has found those arguments persuasive.³⁹

*12 INTRODUCTION

Against a backdrop of recent developments that reveal the gross shortcomings of previously accepted forensic techniques, our empirical evidence and data--derived in large part from our litigation, both nationally⁴⁰ and in Mississippi⁴¹--demonstrates that certain forensic science disciplines are significantly more problematic than previously thought. In fact, their deficiencies are far more egregious than the 2009 National Research Council of the National Academy of Sciences report's ("NAS Report") characterization of them as "imprecise or exaggerated" and the cause of "erroneous or misleading evidence" suggests.⁴² Although *13 there exist scholarly critiques of the generalized shortcomings of this so-called "first generation" forensic evidence⁴³--particularly hair microscopy and bite marks, which are the focus here⁴⁴--our argument relies on empirical evidence, the historical record, and recent scientific and scholarly advancement.

More specifically, our thesis is situated on five bases: (1) the ever-increasing numbers of post-conviction exonerations, particularly those involving bite mark and hair microscopy evidence;⁴⁵ (2) the publication and widespread acceptance of the NAS Report, including recent federal *14 legislative and policy initiatives which call for the realization of several of the Report's core recommendations;⁴⁶ (3) the Federal Bureau of Investigation (FBI) and Department of Justice's (DOJ) unprecedented audit of thousands of hair comparison cases, stemming from its admission that FBI analysts routinely

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE..., 4 Va. J. Crim. L. 1

proffered scientifically invalid testimony in these cases;⁴⁷ (4) state-level legislation amending habeas corpus statutes in order to provide avenues of post-conviction review for petitioners whose convictions rest on discredited scientific **evidence**;⁴⁸ and, (5) our own case data about bite mark identification and hair **evidence**, which provides abundant empirical support for the proposition that ethical and legal obligations should flow as a matter of course with regard to convictions based on such **evidence**.

Several far-reaching consequences logically and inevitably follow from the shifted paradigm thesis advanced here. Among these--each of which we address in turn--are: (a) the lack of scientific or evidentiary validity for certain types of pattern and identification techniques; (b) that, as a result of our empirical findings, state--and to some extent federal--jurisprudence which holds that this type of **evidence** is admissible is objectively erroneous and must be reevaluated and effectively rejected as invalid precedent;⁴⁹ and, (c) that the long-overdue awakening to scientific ***15** reality presents a unique ethical challenge to the profession, one that current professional ethics fail to adequately capture, even though fundamental due process norms compel the conclusion that prosecutors, defense attorneys, experts, and their respective governing bodies have an ethical, moral, and legal duty to revisit convictions resting on discredited scientific **evidence** and provide effective remedies.

Part I sets forth the broad, contextual bases of our argument. Briefly discussed is the predominant incidence of flawed forensic science--specifically bite mark and hair microscopy **evidence**--as a leading cause of **wrongful convictions**. The discussion of the NAS Report's findings in this Part centers on elements of the federal forensic reform agenda that are complementary to or, in certain instances, adoptive of suggestions contained in the Report. Also explored here are recent state efforts to address forensic malfeasance and discredited scientific **evidence** through legislation passed in direct response to post-conviction courts elevating procedural rules over the reality of scientific progress. In Parts II and III, we introduce our own data, as well as its contextual background within **wrongful convictions** generally and the specific disciplines of bite mark identification and hair microscopy science. We likewise provide a diagnosis for the pervasiveness of these forensic disciplines, namely an embarrassingly lax and self-perpetuating approach to the admissibility of unvalidated and false forensic **evidence**. In Part IV, we conclude with a discussion of the unique set of ethical challenges--and pressing obligations--that, left uncorrected, threaten the legitimacy of the justice system.

I. "THE SHIFTED PARADIGM"

Law and science are truth-seeking processes and therefore share a critical, but sometimes anomalous, relationship; "[s]cience helps the law understand the world in which legal policy must operate."⁵⁰ Whereas law values and relies on precedent to establish guarantees of trustworthiness, ***16** scientific inquiry accepts precedent only as a baseline from which to seek a new way forward, sometimes quite rapidly.⁵¹ To the extent consistency and finality--component parts of precedent--are valued in science, it is only insofar as they remain scientifically valid. Put differently, falsified hypotheses are quickly discarded and, if referenced at all, are used to draw a line from what mankind once thought to be true to the current state of scientific knowledge. The law, on the other hand, has moved glacially to abandon techniques which science has proven false or exposed as baseless speculation, with little inquiry.⁵² Exacerbating this problem is the adversarial system's propensity to value zealous advocacy over sound science,⁵³ particularly when deployed against criminal defendants.⁵⁴

The rapid development and introduction of cutting-edge science in our courtrooms has intensified the tension, both with regard to emerging techniques and those with long, if undistinguished, histories. This is particularly true as it relates to traditional forensic individualization sciences. In their 2005 groundbreaking--and controversial⁵⁵--article,

The Coming Paradigm Shift in Forensic Identification,⁵⁶ Michael J. Saks and Jonathan J. Koehler, argued that “[l]egal and scientific forces are *17 converging to drive an emerging skepticism about the claims of the traditional forensic individualization sciences. As a result, these sciences are moving toward a new scientific paradigm.”⁵⁷ Calls for reforming the way the criminal justice system currently views and admits forensic **evidence** in court quickly followed suit.⁵⁸ Among the numerous examples--some of which are novel and progressive--are those that argue that because of the surfeit of documented forensic error,⁵⁹ “suspect *18 evidentiary categories” which “are both recurring features of **wrongful convictions** and not otherwise susceptible to correction through traditional trial mechanisms ... should be subjected to heightened scrutiny for reliability under the Due Process Clause.”⁶⁰ There are also proposals that would subject all “police generated” **evidence**-- namely “eyewitness identification testimony, police officer testimony regarding a defendant's confession, and a police informant's testimony regarding a defendant's incriminating statements”--to a pre-trial reliability screening prior to being offered into **evidence**.⁶¹

Scholars are not the only proponents of reform; courts, too, have recently, albeit belatedly, joined the effort. In 2012, the Supreme Courts of two states--New Jersey and Oregon--issued opinions which redrew the landscape of those state courts' treatment of eyewitness identification **evidence**, a landscape created by the near universally adopted, yet scientifically flawed, “balancing test” announced by the Supreme Court in *Manson v. Brathwaite*.⁶² In *State v. Henderson*,⁶³ the New Jersey Supreme *19 Court assessed decades of social science research regarding the vagaries of eyewitness identification⁶⁴ and, as a result, revised both standards for the **evidence's** admissibility⁶⁵ and the instructions to juries about how to assess the **evidence's** purported value.⁶⁶ In *State v. Lawson*,⁶⁷ the Oregon Supreme Court's decision not only used many of the same factors employed by the *Henderson* court to reverse the conviction and establish a new evidentiary standard for the admission of such **evidence**,⁶⁸ but also went further, shifting the burden of admissibility to the prosecution.⁶⁹ As progressive as these proposed remedies and substantive evidentiary changes are, however, they provide only prospective relief; they are not directed at the continuing failures of justice identified by the data and illustrated by the developments discussed below.

A. FORENSIC **EVIDENCE**-BASED POST-CONVICTION EXONERATIONS

The phenomenon of post-conviction exonerations is now well-known and has been documented elsewhere.⁷⁰ For the purposes of our argument, however, the frequency with which flawed forensic **evidence** serves as a leading cause of **wrongful convictions** is worth reiterating briefly. As a prominent study of the first 200 post-conviction exonerations conducted by Brandon Garrett illustrates, 57% of those convictions *20 involved flawed forensic **evidence**.⁷¹ That statistic is consistent with another figure concerning **wrongful convictions**; 60% of the forensic witnesses who testified in **wrongful conviction** cases provided inaccurate testimony.⁷² More specifically, “[f]orensic **evidence** was the second leading type of **evidence** supporting ... [the first 200] erroneous convictions.”⁷³ Within that subset, serological analysis was the most commonly admitted, followed by hair **evidence**, bite mark **evidence**, and fiber comparison, respectively.⁷⁴ Correspondingly--and as our empirical **evidence** below supports--Garrett's study indicates that these cases not only involved the “use of **evidence** with limited probative value, but [also] the improper use of then-existing forensic science. To a surprising extent, the forensic testimony at trial was improper based on science at the time.”⁷⁵ Hair microscopy testimony, for example, played a role in 22% of the cases which comprised the first 200 exonerations.⁷⁶ Recent investigative reporting at *The Washington Post* revealed something else about this subset of cases, however: that, for years, the DOJ officials who had reviewed work in these cases⁷⁷ and were aware that false or exaggerated testimony provided by its analysts had led to flawed convictions across the country, did not adequately inform defendants *21

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE..., 4 Va. J. Crim. L. 1

whose convictions were affected.⁷⁸ DOJ officials later took the position that the limited notification comported with their legal and constitutional obligations and they “were not required to inform defendants directly.”⁷⁹ Worse, the case review was limited, even though officials were aware that the potential problem was far broader; the FBI not only made its own experts available, but also trained “about 600 examiners from outside the FBI between 1973 and 1987, as well as “an additional 450 examiners ... over the next dozen years.”⁸⁰ The FBI-trained experts were taught to provide the same testimony the FBI has conceded is scientifically invalid.⁸¹

*22 Similarly, a 2013 investigation by the Associated Press (AP) revealed that at least twenty-four innocent men whose convictions and/or indictments were obtained through the use of bitemark **evidence** had been exonerated since 2000.⁸² Consisting of “decades of court records, archives, news reports,” and interviews with “[t]wo dozen forensic scientists and other experts ... including some who had never before spoken to a reporter about their work,” the AP investigation was the “most comprehensive” audit of bite mark cases ever undertaken.⁸³ What is most astounding about the number of **wrongful convictions** based at least in part on bite mark analysis which have been discovered thus far, however, is that the technique is rarely used.⁸⁴

***23 B. NATIONAL ACADEMY OF SCIENCES REPORT AND RECENT COMPLEMENTARY FORENSIC DEVELOPMENTS**

Even acknowledging its critiques,⁸⁵ the NAS Report has been widely recognized as a game-changing document. Central to its concerns were findings associated with the use of flawed forensic science in criminal prosecutions, specifically the “potential danger of giving undue weight to [forensic] **evidence** and testimony derived from imperfect testing and analysis”⁸⁶ For too long, as the Report noted, forensic science has been largely advanced within a legal, rather than a scientific, construct.⁸⁷ Despite the dedicated work of forensic experts, the disaggregated, uncoordinated nature of the system and its legal focus has prevented interested communities from “establish[ing] strong links with a broad base of research universities and the national research community.”⁸⁸ Without the integration of the research community, the forensic science system was deprived of the scientific research funding necessary to meet the foundational and innovational needs; left without measurement and technical standards to guide practice; and, isolated from other scientific communities that have improved many foundational issues (cognitive bias, root cause analyses, laboratory quality, etc.).⁸⁹ The Report went on to conclude that, apart from nuclear DNA analyses,⁹⁰ claims about individualization--including bite mark analysis and hair microscopy, among others--are unsupported by the most basic ***24** foundational research that would allow such claims to be classified as sound science.⁹¹

The scholarship and other reform that the Report has generated has been plentiful, much of it aimed at developing and refining solutions to the problems that the Report identified.⁹² In addition, as discussed below, several practical suggestions for which the NAS Report advocated have been implemented, primarily with federal government impetus and support.⁹³ These developments have also raised the promise of a future for forensic science that engages both the legal and scientific communities. Though this is a new endeavor for both communities, proper support will lead to a successful collaboration of researchers, forensic practitioners, and law enforcement and result in data-driven methods and practicable standard implementation.⁹⁴

1. National Commission on Forensic Science

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE..., 4 Va. J. Crim. L. 1

In 2013, the Department of Justice (DOJ) and the National Institute of Standards and Technology (NIST) signed a Memorandum of Understanding (MOU)⁹⁵ outlining the framework for collaboration in *25 strengthening the validity and reliability of forensic sciences. The MOU provides clear guidance on how the DOJ and NIST will work together to “enhanc[e] oversight and improv [e] coordination across a broad range of forensic science disciplines.”⁹⁶ Among the enhancements is the creation of a new federal advisory committee, the National Commission on Forensic Science (NCFS)⁹⁷ and the creation of discipline-specific guidance groups housed within NIST.⁹⁸

NCFS is charged with providing policy recommendations regarding forensic science to the Attorney General and, more specifically, with “strengthening the validity and reliability of the forensic sciences,” improving quality assurance and quality control in forensic labs, and *26 identifying protocols for **evidence** collection, analysis, and reporting.⁹⁹ A central goal of NCFS is to advise the Attorney General on “the intersection of science and the courtroom” and to recommend standards and policies for implementation at federal law enforcement laboratories.¹⁰⁰ At the first meeting of the NCFS on February 4 and 5, 2014, the Commission members--an impressive array of academic and research scientists, lawyers, judges, forensic science practitioners, and crime lab directors¹⁰¹--began their work by suggesting various subcommittees tasked to specific charges of the MOU.¹⁰²

2. Basic and Applied Research and Standards Development

The NIST also agreed to support the objectives of the Commission through a two-pronged effort. First, NIST “awarded Iowa State University (Ames, Iowa) up to \$20 million over five years to establish a Forensic Science Center of Excellence focused on pattern and digital **evidence**.”¹⁰³ The Forensic Science Center of Excellence will conduct basic foundational research on forensic techniques that, even in the absence of this research, *27 have nevertheless long been accepted by criminal courts.¹⁰⁴ The research “will focus on improving the statistical foundation for fingerprint, firearm, toolmark, dental and other pattern **evidence** analyses, and for computer, video, audio and other digital **evidence** analyses.”¹⁰⁵

Second, NIST will administer and coordinate the Organization of Scientific Area Committees (OSAC).¹⁰⁶ This second prong involves the creation of “a sustainable infrastructure that will produce best practices, guidelines, and standards to improve [the] quality and consistency” of the forensic science disciplines.¹⁰⁷ OSAC's technical standards will augment the measurement standards that are developed independently by the Center for Excellence. OSAC will transition the currently independent Scientific Working Groups (SWGs)¹⁰⁸ into “subcommittees” that will consider their *28 previous work product and engage in new standards-setting activities.¹⁰⁹ The OSAC “will be practice-focused” but “will not provide advice to the Attorney General ... or the NCFS” directly.¹¹⁰ While NIST will administer the OSAC, its membership will be appointed by a NIST/DOJ leadership and membership selection committee.¹¹¹ The NIST Forensic Science Program recently selected 402 experts to serve as members of the 23 subcommittees that make up the five Scientific Area Committees on Biology/DNA, Chemistry/Instrumental Analysis, Crime Scene/Death Investigation, Digital/Multimedia, and Physics/Pattern Interpretation.¹¹²

C. FEDERAL CASE AUDIT & STATE LEGISLATION

The seismic pivot toward the use of validated science in criminal prosecutions discussed in the preceding sections is forward-looking reform. At least as pressing an issue emerging in this new era of scientific integrity is how the criminal justice system addresses convictions resting on discredited expert testimony. To that end, in what should be the beginning

of a series of complementary efforts,¹¹³ the FBI and DOJ *29 recently announced an unprecedented audit--both in mission and scope--of all FBI Laboratory hair comparison cases since the early 1980s.¹¹⁴ Questions about the validity of the FBI's training of its forensic examiners--including the training it provided about how those examiners should testify about their purported findings--came under intense scrutiny after two cases in Washington, D.C. were the subject of post-conviction exonerations.¹¹⁵ In 1980, Santae Tribble was charged with and convicted of the murder of a D.C. taxi driver.¹¹⁶ Tribble's conviction was based almost entirely on an FBI agent's testimony that hair discovered in a stocking mask "matched [Tribble's hair] in all microscopic characteristics."¹¹⁷ In December of 2012, Tribble was exonerated and released from prison after DNA testing excluded Tribble as the source of the hair.¹¹⁸ Tribble's exoneration had been preceded by Kirk Odom's, *30 another D.C. defendant who had been convicted of a 1981 rape.¹¹⁹ Odom served more than twenty years in prison.¹²⁰ At Odom's trial, an FBI analyst testified that a hair found on the victim's nightgown was "microscopically like" Odom's; yet, even according to the prosecution, the analyst had only "been able to distinguish between hair samples ... 'eight or 10 times in the past 10 years, while performing thousands of analyses.'"¹²¹ The FBI and the DOJ have acknowledged that this type of testimony is scientifically invalid, along with two other varieties of testimony FBI examiners routinely proffered to jurors in an effort to quantify the significance of an association between a questioned and known hair.¹²²

*31 The Tribble and Odom cases were unique in their influence, but the scientifically invalid testimony used to secure their convictions was, as discussed below, routine and widespread. Aside from the false and misleading expert testimony and the failures of justice that resulted, at least as troubling is the initial response to the problem by the agencies responsible for it. As *The Washington Post* journalist Spencer Hsu documented in a series of articles, federal officials began reviewing these types of cases¹²³ in the 1990s. But, instead of releasing this information--information which demonstrated that the testimony and the forensic work on which it was based was flawed--to the defendants whose convictions were affected, the federal task force only made it available to prosecutors.¹²⁴

Nonetheless, the audit has already resulted in several firsts. Recently, the *Post* reported that no fewer than twenty-seven death penalty *32 convictions from around the country are among the affected cases.¹²⁵ Among them is the case of Willie Manning, who had been scheduled for execution in Mississippi in May 2013 for the abduction and murder of two college students.¹²⁶ The Mississippi Supreme Court denied Manning's request for post-conviction DNA testing the week prior to his scheduled execution.¹²⁷ In the days immediately preceding his execution date, however, the FBI and DOJ jointly wrote letters to Manning's counsel and Mississippi officials explaining that Manning's case had been included in the audit because the testimony provided by the FBI analyst in the case--that hair found at the crime scene implicated Manning--was "erroneous" and "exceeded the limits of the science" by claiming that the analysis could match the hair to an individual with "a relatively high degree of certainty."¹²⁸ Only hours before he was to die, the Court granted a stay *33 that was presumably based on the audit's findings.¹²⁹ Later that summer, the Court granted Manning leave to seek post-conviction DNA testing.¹³⁰

Similar retrospective efforts to identify and correct tainted convictions have gained ground in Texas and California. In the wake of appalling revelations about forensic malfeasance in criminal trials, executive meddling in efforts to right wrongs, and indefensible¹³¹ and inconsistent¹³² legal rulings, both states enacted legislation aimed at *34 correcting the wrongs. In 2005, Texas created its Forensic Science Commission,¹³³ whose mission is to "investigate, in a timely manner, any allegation of professional negligence or misconduct that would substantially affect the integrity of the

results of a forensic analysis conducted by an accredited laboratory, facility, or entity.”¹³⁴ In 2009, however, on the eve of the Commission's report on Cameron Todd Willingham's case¹³⁵--which was expected to find that the fire science used to secure Willingham's conviction and death sentence was fundamentally flawed¹³⁶--Governor Rick Perry declined to re-appoint the Commission's chairman and appointed three new commissioners.¹³⁷ The Willingham report was delayed.¹³⁸ In 2013, the Texas Legislature enacted several reforms, all aimed to one degree or another at the State's stunning incidence of **wrongful convictions**, many of them based on flawed science.¹³⁹ Perhaps foremost among the efforts is Senate Bill 344, ***35** which allows challenges to convictions gained through now-discredited forensic techniques.¹⁴⁰

Similarly motivated by forensic embarrassments--in this instance a 1997 murder conviction based on bite mark **evidence** and a State Supreme Court decision affirming the conviction that was voted the state's worst appellate decision of the year--California recently passed Senate Bill 1058.¹⁴¹ The legislation allows habeas petitioners to seek relief when a forensic expert repudiates his trial testimony or where that testimony is undermined by scientific or technological advancements subsequent to trial.¹⁴²

***36 D. EMPIRICAL DATA**

Finally, we rest our argument on separate sets of data that deconstruct foundational legal doctrine and demonstrate empirically--when viewed in context with the developments above--the fallacies of the various forensic disciplines, as well as analysts' willingness to tout and courts' willingness to embrace such **evidence** in the absence of basic validation research. This phenomenon has been noted elsewhere, though its prevalence can now be definitively illustrated after several years of exonerations involving bite mark and hair microscopy testimony. For example, in another leading article, data illustrates that courts “policed the introduction of forensic testimony in these trials in a highly deferential manner, typically trusting the jury to assess the expert testimony.”¹⁴³ Defense attorneys' failure to challenge the **evidence** in the first instance--combined with appellate courts' failure to take the issues seriously¹⁴⁴--meant that many exonerees never challenged the forensic **evidence**; moreover, even where challenges were raised, courts invoked the ***37** “harmless error” doctrine to avoid granting relief, occasionally affirming plainly fabricated statistical conclusions.¹⁴⁵ In those instances where exonerees did challenge the putative scientific **evidence**, appellate courts typically relied on decades--and sometimes centuries--of precedent supporting the admissibility of the technique at issue to brush aside requests for relief.¹⁴⁶

We have termed this phenomenon the “Echo Chamber.” Courts fail to engage in a meaningful review of the proffered **evidence** through either a *Frye*¹⁴⁷ or *Daubert*¹⁴⁸ hearing and, instead, cite “persuasive” authority from sister states admitting such **evidence**, even in cases of first impression.¹⁴⁹ In other cases, courts admit the technique based on some other rationale, typically that analysts--often those testifying, who have a professional interest in the technique's continued admissibility--agree that the **evidence** at issue is “generally accepted” within their own “scientific community.”¹⁵⁰ A third line of reasoning leading to the uncritical admission of invalid scientific **evidence** involves abdicating judicial gatekeeping responsibly entirely and allowing juries to evaluate competing opinions, or even the legitimacy of the discipline itself.¹⁵¹

***38 II. BITE MARK EVIDENCE**

A. INTRODUCTION

Perhaps no discredited forensic assay has benefited more from criminal courts' abdication of gatekeeper responsibilities than bite mark analysis. The genesis of the flawed jurisprudence can be traced back to a single case: *People v. Marx*,¹⁵² the first reported case to consider the admissibility of bite mark comparison **evidence** in human flesh.¹⁵³ Although there was no pre-trial *Frye*¹⁵⁴ hearing--or any other **evidence** or rule-based admissibility hearing--to examine the validity and reliability of the new technique in *Marx*,¹⁵⁵ over time, *Marx* has proven to be a seminal decision. *Marx* turned an obscure, unvalidated sub-discipline of "forensic odontology"¹⁵⁶ into mainstream, "generally accepted," "scientific" **evidence**¹⁵⁷ without any basic or applied research to validate the technique's two underlying hypotheses¹⁵⁸ that: (1) a properly trained *39 bite mark expert can make an association between a bite mark and a suspect's "dentition" (the biting surface of teeth); and, (2) that a properly trained expert can provide a scientifically valid estimate of the rareness or frequency of that association (i.e., how many other dentitions may also be *40 associated with the putative bite mark).¹⁵⁹ Despite criminal courts' continued acceptance of bite mark analysis, no such validation research exists today.¹⁶⁰

Marx's influence demonstrates the disinclination of criminal courts to engage in a rigorous analysis of putative scientific **evidence**--at least in criminal cases¹⁶¹--or even to scrutinize precedent carefully when *41 weighing the admission of such **evidence**. Indeed, the *Marx* court clearly recognized that basic tenets of science¹⁶²--generating a hypothesis, testing that hypothesis through laboratory and field experiments, publishing the results in peer-reviewed journals, repeating the experiments, and testing the results of those experiments under a wide variety of conditions--were entirely absent from the nascent field. The court itself wrote that there was "no established science of identifying persons from bite marks";¹⁶³ that "experts do not agree on the exact number of similarities necessary to make a positive identification;"¹⁶⁴ and that "there was no **evidence** of systematic, orderly experimentation in the area."¹⁶⁵ Nonetheless, the *Marx* court found that, "[l]eaving aside the question whether tooth bites made into human flesh are sufficiently common in forensic dentistry to expect that orderly experimentation will ever be possible,"¹⁶⁶ the bite mark identification testimony was admissible. According to the court, the standard of "general acceptance"¹⁶⁷ by recognized experts--i.e., the *Frye* test--was *not* determinative of admissibility because "the basic data on *42 which the experts based their conclusion was verifiable by the court."¹⁶⁸ The court found that because it was able to observe with its own eyes--*Marx* was a bench trial--the "matching" of the defendant's dentition with the bite mark at issue it need not "sacrifice its ... common sense in evaluating it"¹⁶⁹ and could independently verify the conclusions the experts were urging. "Indeed," the court wrote, "it is evident that in most cases the expert himself must accept certain dogmas of his profession on faith. We doubt that the average criminologist could supply the data on which the reliability of fingerprint **evidence** is based."¹⁷⁰ Thus was laid the foundation for the admissibility of bite mark analysis. Over the past three decades it has served as the foundation of at least twenty-four **wrongful convictions** and indictments.¹⁷¹

B. THE RISE OF BITE MARK MATCHING

Exacerbating the error is the fact that subsequent cases of *first impression* in other states relied on *Marx* as precedent for both the proposition that a *Frye* hearing is unnecessary prior to the admission of bite mark **evidence** and the admissibility of bite mark analysis generally as "scientific" **evidence** for everything from dog bites¹⁷² to bites made in paper towels.¹⁷³ As the graphic below demonstrates, like *Marx*, the *43 overwhelming majority of these cases fail to examine the

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE..., 4 Va. J. Crim. L. 1

reliability of the technique prior to its admission at trial. Unlike *Marx*, however, all but three decisions arbitrarily label the **evidence** as “scientific.”¹⁷⁴ For example, the next post-*Marx* California court to consider bite mark **evidence**, in citing to *Marx*,¹⁷⁵ noted the “superior trustworthiness of the scientific bite mark approach,”¹⁷⁶ which compared favorably to “other scientific-test **evidence**,”¹⁷⁷ such as the “breathalyzer test.”¹⁷⁸

TABULAR OR GRAPHIC MATERIAL SET FORTH AT THIS POINT IS NOT DISPLAYABLE

¹⁷⁹

*44 The three reported cases of first impression in which trial courts actually held *Frye* hearings demonstrate another deep flaw running *45 through the jurisprudence: the failure to distinguish between the methodology experts employ to collect data and the scientific basis for interpreting the data collected.¹⁸⁰ The former is typically well established, non-controversial, and impressive to triers of fact; the latter is often entirely absent.

Bite mark experts--who refer to themselves as “forensic odontologists”-- employ an imposing array of techniques and analytical instruments to collect data and analyze teeth and the bite marks they purportedly leave.¹⁸¹ For example, the photographic techniques alone include infrared,¹⁸² ultraviolet,¹⁸³ and trans-illumination photography,¹⁸⁴ which penetrates below the surface of the skin. To ensure photographs of bite marks are to scale, an “American Board of Forensic Odontology Number 2 Photomacrographic Scale”¹⁸⁵ --a right angled ruler--is placed on the same plane as the alleged bite mark. Molds of suspect dentitions are *46 created using a substance known as “polyvinylsiloxane.”¹⁸⁶ Finally, “digital overlays”¹⁸⁷ or outlines of exemplar “dentitions”-- the biting surface of teeth--are generated through a digital software program and then used to superimpose the dentition to the photograph of the bite mark.¹⁸⁸ Forensic odontologists also make liberal use of scientific jargon and technical dentistry terms, such as the “occlusal view” of teeth (closed mouth);¹⁸⁹ “avulsive” bite marks (flesh tearing);¹⁹⁰ “central ecchymosis” in the bite mark (bruising in the center of a bite mark);¹⁹¹ “subcutaneous hemorrhage” (bleeding beneath the skin);¹⁹² “temporomandibular joint function” (an individual's biting mechanics);¹⁹³ anterior and maxillary dentition (upper and lower teeth);¹⁹⁴ and other esoteric terminology.

Courts considering admissibility challenges--and jurors weighing life and liberty issues--are impressed by the facility these experts have with the language of science and the supposed precision and complexity of the data collection process.¹⁹⁵ However, while the terminology used by *47 forensic odontologists is beyond the ken of a typical lay juror, the core methodology is easily understood. In essence, experts place outlines of suspects' teeth over photographs of bite marks and decide whether they “match.” That the analysis of the data is entirely subjective and lacks any basis in science is typically lost on courts and factfinders. The few reported *Frye* hearings reflect this misunderstanding.¹⁹⁶

Wisconsin's case of first impression, *State v. Stinson*,¹⁹⁷ is a paradigmatic example. *Stinson* is one of three cases in which a pre-trial admissibility hearing was held before *State v. Armstrong*¹⁹⁸ became the first case to take judicial notice of the supposed “general acceptance” of bite mark **evidence** in the scientific community.¹⁹⁹ Ruling that the trial court did not abuse its discretion in finding that “there are adequate standards and controls in the area of forensic odontology, specifically for the identification of an individual through bitemark [sic] **evidence** and that *48 that area of science is an accepted area ... of science,”²⁰⁰ the Wisconsin Court of Appeals did not evaluate either of the two fundamental hypotheses of bite mark analysis, namely that: (1) a properly trained analyst can make that type of association; and, (2)

that the analyst can provide a scientifically valid estimate of that association's evidentiary value.²⁰¹ Instead, it focused on the impressive credentials of the experts and what appeared to be their sophisticated data collection and comparison techniques. The following excerpt from the opinion is worth quoting at length:

Dr. Lowell Thomas Johnson ... a practicing dentist and a clinical professor of pathology at Marquette University School of Dentistry, testified for the state. On November 3, 1984, Dr. Johnson was called by the medical examiner and asked to examine the [63-year-old] victim's body. Upon examining Cychosz [the victim], Dr. Johnson discovered eight complete or partial bite marks. To preserve this **evidence**, Dr. Johnson had a photographer from the state crime laboratory photograph the bite marks. Dr. Johnson then made a rubber impression of the victim's right breast which (sic) contained the greatest number of three-dimensional indentations. According to Dr. Johnson, when the wounds are three-dimensional, or when there are any indentations present, they can be well preserved by taking an impression of them. This impression is then later used to produce a static replica of the bite marks which (sic) will not be subject to distortion.

***49** Dr. Johnson also testified that as part of established procedure, he preserved some of the tissue from the deeper bites. This was done by affixing an acrylic ring to the tissue surrounding the indentations and then removing that block of tissue for future study.

In addition to examining Cychosz, Dr. Johnson also did a complete forensic workup on Stinson. As part of the workup, a special camera was used to photograph the biting and facial surfaces of the teeth. A set of rubber impressions were then taken so a model of Stinson's teeth could be prepared. In addition, Dr. Johnson examined Stinson's teeth to observe the presence of defective or decayed teeth, or teeth which had been artificially restored.

Dr. Johnson also performed a similar dental workup on Robert Earl Stinson, the defendant's twin brother. Based on his comparison of the **evidence** taken from the victim with the models of Robert Earl's teeth, Dr. Johnson concluded that there were some gross discrepancies which would rule out Robert Earl Stinson as having possibly made the bite marks.²⁰²

Dr. Johnson next testified extensively on the comparisons he made using the dental impressions of Stinson's teeth and the bite marks found on the victim's body. He described and demonstrated the methods he used in making these comparisons. First, a comparison was made using the model of the bite marks and the model of Stinson's teeth. A comparison was also made by placing the model of Stinson's teeth over photos of the bite marks to ***50** see if the features were consistent. In addition, Dr. Johnson used an overlay technique, which he stated was another standard procedure in bite mark comparison. By taping a black and white negative of Stinson's teeth over a color transparency of the bite mark, Dr. Johnson was able to compare the patterns of the bite marks with the patterns of the teeth. Based on these comparisons, Dr. Johnson concluded that the bites he examined on Cychosz "had to have been made by teeth identical in all of these characteristics to those that I examined on Robert Lee [Stinson]."

The state also called Dr. Raymond Rawson (Dr. Rawson), a forensic odontologist, who, as chairman of the Bite Mark Standards Committee of the American Board of Forensic Odontologists, participated in formulating the standards and procedures for evaluating bite mark **evidence**. Dr. Rawson was asked to conduct an independent evaluation of the bite mark **evidence** using Dr. Johnson's models and photos. Dr. Rawson testified that the **evidence** in the case was “high quality” and “overwhelming.” He stated that this was an “exceptional” case because “[t]here were more ... pieces of **evidence** than you usually see in a bite mark case.”

After examining Dr. Johnson's workup, Dr. Rawson stated that the methods Dr. Johnson used in gathering the **evidence** complied with the standards of the American Board of Forensic Odontology. Dr. Rawson then analyzed the **evidence** and concluded, to a reasonable degree of scientific certainty, that Stinson had inflicted the bite marks found on [the victim's] body.

Dr. Rawson also reviewed the **evidence** produced from the examination of Stinson's twin brother. Dr. Rawson *51 testified that after an extensive analysis of the similarities and differences between the two brothers' mouths, he found significant discrepancies in their dentition. Therefore, Dr. Rawson concluded, Robert Earl Stinson could not have inflicted the bite marks found on Cychosz's body

Dr. Johnson stated that the availability of bite marks from different parts of the body eliminated the possibility that the impressions obtained may have been distorted. He also testified as to the methods used in preserving and comparing the bite mark **evidence** gathered.

A total of fourteen upper and lower jaw impressions were made from the bite marks found on Cychosz's body. Because of the opportunity to examine so many bites, and the fact that some of the bites were so deep as to be three-dimensional, Dr. Johnson testified he was able to detect a repetition of some particularly unique features in several of the bites.

Dr. Johnson later performed a forensic odontological examination of Stinson. Following the examination, Dr. Johnson noted the following unique features: one of the central incisors was fractured and decayed almost to the gum line; the lateral incisor in the upper jaw was set back from the other teeth; all of the upper front teeth were flared; the lower right lateral incisor was worn to a pointed edge; the right incisor was set out from the other teeth on the lower jaw. Dr. Johnson used these features along with the arch of the mouth and the spacing, width, and alignment of the teeth to make comparisons with the bite marks found on the victim. After an exhaustive examination of the photos, models and tissue samples taken from Stinson and the victim, Dr. Johnson *52 concluded, to a reasonable degree of scientific certainty, that the bite marks on the victim were made by Stinson.

The jury also heard from Dr. Rawson who concluded, based on the workup Dr. Johnson performed on both the victim and Stinson, that Stinson had inflicted the bite marks on the victim. In Dr. Rawson's opinion the **evidence** in the case was overwhelming and he stated that "if we have four or five teeth that we are able to examine, then we can say that there is no other set of dentition like that." In this case, Dr. Johnson was able to identify seventy-five individual tooth marks in various combinations of between five and eleven teeth.

Based upon this **evidence**, we hold that a jury could reasonably conclude beyond a reasonable doubt that Stinson murdered Cychosz. *The reliability of the bite mark **evidence** in this case was sufficient to exclude to a moral certainty every reasonable hypothesis of innocence.*²⁰³

Stinson was, of course, innocent.²⁰⁴ Although the bite mark **evidence** was presented to the jury as "scientific" **evidence**, the *Stinson* court found that, by the time Stinson's appeal was heard, the state had rejected the *Frye* test--only to have the test reemerge two years later²⁰⁵--and thus that "**evidence** given by a qualified expert is admissible *irrespective of the underlying scientific theory.*"²⁰⁶ Borrowing the reasoning of *Marx*, the court found that bite mark **evidence** was a valuable aid to the jury because "[b]y looking directly at the physical **evidence** used, the models and the photos, the jury was able to judge for itself *53 whether Stinson's teeth did in fact match the bite marks found on the victim's body."²⁰⁷ As **evidenced** by the dozens of **wrongful convictions**,²⁰⁸ jurors are not capable of determining whether bite marks "in fact" match a defendant's dentition or--more generally, but perhaps more importantly--to recognize that some prosecutors are willing to proffer pseudo-science as **evidence** of culpability in lieu of careful and thorough law enforcement investigations.²⁰⁹ Had Stinson's jury relied on its own observations, Stinson may not have spent over two decades in prison. As the federal court adjudicating Stinson's civil suit against the experts who testified at his trial noted, the "eyeball test" showed that his dentition did not "match" the bite mark; he was missing a tooth where the perpetrator appeared to have had one, and there was no explanation as to "why a bite mark was on [the victim's] body where Stinson has a missing tooth."²¹⁰

Stinson's jury of course relied on the interpretation of the data urged by the "scientists," even though there was--and is--no **evidence** that bite mark experts are capable of reliably associating a dentition with a bite mark; experts are not required to and do not undergo proficiency testing.²¹¹ Moreover, even if such associations could be made, there was--*54 and is--no basis in science for concluding that a suspect is the source of the bite mark to the exclusion of all other potential sources.²¹² Put differently, since the two hypotheses underlying bite mark analysis have never been scientifically validated, conclusions offered by these experts are not helpful to a trier of fact because there is no **evidence** the technique is capable of providing probative **evidence**.

Jurisdictions that adopted the *Marx* "eyeball test" allowed proponents of bite mark analysis to have their cake and eat it too. Self-validating experts routinely proffered to juries "scientific" **evidence** of culpability, yet the empirical basis for the conclusions were not required to undergo *Frye* scrutiny,²¹³ or, later, challenges pursuant to *Daubert*²¹⁴ or *56 *Kumho Tire*.²¹⁵ Some states adopted the "eyeball test" specifically to exempt certain pattern-matching forensic techniques from judicial scrutiny, relying instead on lay jurors to distinguish between valid science and subjective speculation masquerading as scientific **evidence**.²¹⁶

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE..., 4 Va. J. Crim. L. 1

Connecticut jurisprudence in this area demonstrates how these pattern-matching techniques were allowed into court through the backdoor *Marx* created for bite mark **evidence**.²¹⁷ The Connecticut Supreme Court adopted *Daubert* as the standard for the admissibility of scientific **evidence** in 1997.²¹⁸ In doing so, the court emphasized the importance of the trial court acting as a “gatekeeper” and assuming responsibility for determining the validity and reliability of scientific **evidence**, noting that “a judge is in a much better position than a juror to assess accurately the *57 fundamental validity of [scientific] **evidence**.”²¹⁹ The court acknowledged that a juror's understanding of scientific **evidence** is “largely dependent on the presentations of the parties and their experts;”²²⁰ that “expert presentations may often be misleading;”²²¹ and, that “cross-examination may often be difficult and ineffective in bringing out flaws in the expert's reasoning.”²²² Judges, on the other hand, “have the benefit of reviewing briefs and other documents”²²³ and demanding “supplemental briefing on any issue that needs clarification.”²²⁴ The court went on to note that “certain types of **evidence**, although ostensibly rooted in scientific principles and presented by expert witnesses with scientific training, are not ‘scientific’ for the purposes of [the] admissibility standard for scientific **evidence**.”²²⁵ Such **evidence** “simply require[s] the jurors to use their own powers of observation and comparison,” and thus does not require a *Daubert* hearing.²²⁶

By the time the West Virginia Supreme Court became the first state high court to take judicial notice of the general acceptance of bite mark **evidence**, twenty-one states had already decided it was admissible, without a single dissenting opinion.²²⁷ Though the jurisprudence does not *58 withstand contemporary scrutiny, it supported--and continues to support--the argument that every state which considered the admissibility issue decided that bite mark analysis passed evidentiary muster, however little muster was required. Subsequent cases of first impression became foregone conclusions. As the chart below demonstrates, courts began citing to one another as a matter of course, creating an echo chamber of ill-considered opinions.

TABULAR OR GRAPHIC MATERIAL SET FORTH AT THIS POINT IS NOT DISPLAYABLE

As noted above, *Marx* looked to *Frye* to support the conclusion that bite mark analysis was exempt from *Frye* scrutiny.²²⁸ This anomalous, yet remarkably influential, reasoning had another pernicious effect on trace **evidence** jurisprudence: allowing the experts themselves to define the “relevant scientific community.” In *Frye*--which involved a challenge to the admissibility of a lie detector test--the court precluded the **evidence** *59 because it had not yet “gained such standing and scientific recognition among physiological and psychological authorities as would justify the courts in admitting expert testimony deduced from the discovery, development, and experiments thus far made.”²²⁹ Had the *Frye* court defined the scientific community as lie detector practitioners, the **evidence** would have been admissible, since it presumably had achieved “scientific recognition” within that narrow community.

Precedent-establishing bite mark analysis cases across the country, however, uniformly defined the relevant scientific community as the forensic dentists themselves.²³⁰ Those who “generally accepted” the discipline--many of whom practiced in the field and had a vested interest in its success--constituted the relevant community whose general acceptance was required for admissibility.²³¹ The self-referential and self-interested community essentially resulted in the question of the field's admissibility being a foregone conclusion. By the time that the NAS *60 Committee--which was comprised of leading scientists in all relevant fields²³²--examined the validity and reliability of bite mark analysis and concluded that the technique was inherently weak and lacked any basis in science in 2009, decades of state court precedent had reached exactly the opposite conclusion.²³³

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE..., 4 Va. J. Crim. L. 1

Finally, it is worth noting, albeit briefly for purposes of substantive discussion, that a research team led by Dr. Mary Bush--a tenured professor at the State University of New York at Buffalo's School of Dental Medicine and past president of the American Society of Forensic Odontology--began to develop studies in the field subsequent to the publication of the NAS Report and the attention the Report brought to the shortcomings of forensic odontology.²³⁴ Twelve studies that tested the foundational issues related to skin's ability to act as a substrate for interpreting data were ultimately conducted.²³⁵ Each used a cadaver model, and each was published in a peer-reviewed scientific journal.²³⁶

*61 Broadly speaking, the studies' research strongly suggests what is intuitive; even assuming the uniqueness of human dentition, human skin is not capable of capturing that uniqueness with sufficient fidelity to identify "the biter."²³⁷ Moreover, bite marks created by the same dentition on the same individual appeared substantially different, depending on the angle of the body and whether the mark was made parallel or perpendicular to "Langer lines."²³⁸ Likewise, in a study conducted using orthodontically *62 treated dentitions--i.e., teeth straightened through orthodontic work-- Dr. Bush and her team found that bite marks created by treated dentitions could not be reliably distinguished from each other.²³⁹ More specifically, the research team found that:

[a]s may be anticipated, orthodontic treatment had a very strong effect on dental shape similarity. The match rate in the known orthodontically treated set was 42.7% of individuals using the same threshold parameter in only 110 specimens. This confirmed that when orthodontically treated or naturally well-aligned, dentitions may be indistinguishable. This result is also a measure of how successful orthodontic treatment is at producing homologous dental arch shapes. The orthodontically treated human dentition is not unique, as measured here with high accuracy and precision.²⁴⁰

The Bush studies confirm the NAS Report's observation that the "validity of forensic odontology" may be "severely limited" because it relies on interpreting data from a bite mark, which "will change over time and can be distorted by the elasticity of the skin, the unevenness of the surface bite, and swelling and healing."²⁴¹ It is for this same reason that there are no measurement processes or objective standards for bite mark analysis.²⁴² The field simply has no methodology to account for the great variation in the size and shape of the bite marks created by the same *63 dentition.²⁴³ Moreover, manipulating a mold of a suspect's teeth on the victim's decomposing body and declaring a "match" is plainly a scientifically invalid method, incapable even of associating a particular dentition with a bite mark,²⁴⁴ to say nothing of the false individualization claims made in so many cases.²⁴⁵

III. HAIR EVIDENCE: UNCHALLENGED AND UNVALIDATED

A. INTRODUCTION

Microscopic hair comparison attempts to link a known hair--typically from a suspect in a criminal case--and a questioned hair, typically a hair found at a crime scene.²⁴⁶ Like bite mark analysis--indeed all pattern and impression forensic assays²⁴⁷--this technique rests on two hypotheses: (1) that a properly trained hair examiner can make an association between a questioned sample hair and sample hair from a suspect; and, (2) that a properly trained hair examiner can provide a scientifically valid estimate of the rareness or frequency of that association. As discussed *infra*, court-sanctioned, yet scientifically invalid conclusions concerning the second hypothesis have had a long and ignoble history in the United States, and no court has ever rejected the validity of the first hypothesis in a reported opinion. This, despite

the staggering number of **wrongful convictions** obtained at least in part through the use of ***64** hair comparison **evidence** and recent, highly publicized research conducted by FBI hair comparison experts which demonstrates unacceptably high error rates in their own work.²⁴⁸

Judicial acceptance of the first hypothesis has humble origins, not as expert testimony, but as **evidence** introduced through lay witnesses using their own powers of observation to compare known and crime scene hairs. The first reported use of such **evidence** may be traced back to a murder investigation on a cotton plantation in Sunflower County, Mississippi, where, in 1855, John Browning and his son, Gaston Browning, were tried for the murder of John Neal, the overseer of the Hill and McNeill Plantation.²⁴⁹ Amongst other injuries, Neal's neck had been dislocated and broken.²⁵⁰ A search of the defendants' home revealed a noose with drops of blood; moreover, “[o]n the rope near the noose were found several hairs, which upon comparison corresponded exactly in color and length with Neal's hair.”²⁵¹ Despite this **evidence**--and a not insignificant amount of additional circumstantial **evidence**--one justice of the Mississippi Supreme Court found the **evidence** insufficient to sustain the capital conviction, and neither defendant was ever convicted of the crime.²⁵²

***65** Thirteen years later, in *Commonwealth v. Dorsey*,²⁵³ the Supreme Judicial Court of Massachusetts approved the introduction of hair comparison testimony by two lay witnesses, one of whom claimed that hairs found on a club alleged to have been the murder weapon appeared to be human hairs and the other of whom claimed that the hairs “resembled the hair of the deceased.”²⁵⁴ The *Dorsey* court found that observational **evidence** “gained through [the] senses” by lay witnesses was admissible and did not require expert testimony.²⁵⁵ The court, in other words, articulated one of the earliest “eyeball” tests to admit lay opinion testimony of trace **evidence**. Lay testimony concerning handwriting, shoe impression, hair comparison, and **evidence** that a “fragment of a garment” originated from a particular source-- e.g., fiber analysis--was therefore admissible.²⁵⁶ However, “[w]hen other tests than the senses are to be applied to these subjects in order to gain knowledge that cannot be gained by common observation, but must be acquired by the application of special skill or learning, the testimony of experts must be resorted to.”²⁵⁷

Over the next century, trace **evidence** analysis--including hair comparison-- became the province of experts as crime solvers began using increasingly sophisticated instruments to amplify observations of hairs.²⁵⁸ ***66** Although the observations experts made using these instruments--the collection of the data--did indeed require special skill and training, the conclusions concerning “matches” between known and evidentiary hairs--the interpretation of the data, i.e., *the science*--is no more probative today than it was at the turn of the nineteenth century.²⁵⁹ Nor have the core claims of the experts changed since that time.²⁶⁰ As with bite mark analysis, however, courtroom presentations have become more robust. They now often feature elaborate demonstrations accompanied by detailed visual demonstrations of microscopic images, experts referring to themselves as “scientists,”²⁶¹ and the employment of terms such as “reasonable scientific certainty,”²⁶² all without a judicial inquiry into reliability²⁶³ to frame and support the conclusions. The rationale for ***68** abdicating gatekeeping responsibilities has its roots in *Dorsey* and other early jurisprudence in this area.

An examination of these early opinions reveals a persistent line of reasoning that mirrors *Marx's* bite mark analysis: the judicial determination that a comparison of two items to decide whether or not they “match”--even when such analysis is performed by a “scientist” using sophisticated analytical instruments--is not subject to *Frye* scrutiny because conclusions drawn from a comparative examination are elemental, independently verifiable by the trier of fact, and unlikely to mislead or confuse jurors.²⁶⁴ Put differently, the “eyeball test” articulated in *Dorsey* for the admissibility of lay observations of trace **evidence** was eventually extended to shield expert testimony from judicial scrutiny as well. It must be emphasized that the judicial labeling of this or other pattern-matching techniques as “non-scientific” or

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE..., 4 Va. J. Crim. L. 1

“technical knowledge” based on the perception that they are straightforward forensic assays--unlike, for example, DNA genotyping--is not a defensible position. The Supreme Court in *Kumho Tire* rejected the distinction between scientific and technical **evidence** for purposes of applying the *Daubert* test, noting *69 that: “[I]t would prove difficult, if not impossible, for judges to administer evidentiary rules under which a gatekeeping obligation depended upon a distinction between ‘scientific’ knowledge and ‘technical’ or ‘other specialized’ knowledge. There is no clear line that divides the one from the others.”²⁶⁵ Thus *Kumho Tire* eliminated the arbitrary labels that effectively precluded inquiry into the foundational bases of expert testimony considered non-scientific for admissibility issues only. Virtually no state, however, has applied *Kumho Tire* to pattern-matching techniques, even in jurisdictions that have generally adopted *Daubert*.²⁶⁶ Instead, courts have shielded these assays from *Frye/Daubert* scrutiny, while allowing experts to refer to themselves as scientists and exploit the aura of “mythic infallibility” scientific **evidence** often holds over lay jurors.²⁶⁷

Furthermore, inquiries into whether two evidentiary items are indistinguishably similar (i.e., match) and conclusions about how probative those associations are the domain of science.²⁶⁸ More precisely, *70 validating the accuracy of the expert's conclusion is an inherently scientific inquiry. For such an inquiry to result in a scientifically valid conclusion, the interpretation of the data must be drawn from a reliable foundation. Where conclusions by pattern-matching experts are made without objective thresholds for declaring an association or underlying statistical data by which to judge the relevance of the association,²⁶⁹ the conclusions are entirely subjective and therefore unscientific.²⁷⁰ It does not *71 necessarily follow that analyses based on an expert's training and experience are without value,²⁷¹ but the limited value of conclusions drawn from experiential foundation must be plainly communicated to jurors. Expert witnesses must be not permitted to claim the mantle of “science” in courtroom testimony, while at the same time avoiding the judicial scrutiny scientific **evidence** is required to undergo before it is proffered as such to lay jurors.²⁷²

B. THE RISE OF HAIR MICROSCOPY

Apart from some notable early exceptions,²⁷³ the only consistent limits courts placed on hair **evidence** were restrictions on conclusive individualization claims--i.e., explicit assertions that the defendant was the source of the questioned hair to exclusion of all other potential sources--and the use of statistics²⁷⁴ to suggest that a defendant was the probable source of a crime scene hair.²⁷⁵ Although the establishment of the *72 FBI crime lab in 1942 professionalized the technique,²⁷⁶ the result was not *73 better science, but widespread, exaggerated claims of the capabilities of hair **evidence**.²⁷⁷ As discussed earlier, there is **evidence** that the FBI understood the inherent limitations of hair microscopy **evidence** but deliberately obfuscated these limitations through testimony that either implicitly or explicitly argued that association between a suspect's hair and a crime scene hair was highly probative **evidence** of the defendant's presence at the scene.²⁷⁸ Instead of utilizing a validated measurement process with objective thresholds for declaring a “match” (hypothesis one), experts began making claims that the internal characteristics of hair--as observed through the high-powered, side-by-side microscopic examination of questioned and known hairs--revealed similarities of such minute quality that simply “matching” the two hairs was probative **evidence** of the source of the questioned hair.²⁷⁹ Such claims were made *74 despite the fact that there is microscopic variability between hairs originating from the head of the *same individual*.²⁸⁰

*75 The juxtaposed photos below illustrate the persuasive power of visual demonstrations of so-called “matches.” One photo depicts a hair from Jimmy Ray Bromgard; the other a hair recovered during the investigation of the sexual assault of an eight-year-old girl for which Bromgard stood accused.²⁸¹ The FBI-trained hair comparison expert who testified in the prosecution of Bromgard claimed that his analysis revealed that the recovered hair belonged to Bromgard.²⁸²

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE..., 4 Va. J. Crim. L. 1

More specifically, he asserted that the head and pubic hairs found on the sheets were indistinguishable from Bromgard's and that there was less than a one in ten thousand (1/10,000) chance that the hairs did not belong to him.²⁸³ The photo below was used to display the visual similarities. Bromgard was convicted and spent nearly 15 years in prison.²⁸⁴ He was exonerated in 2002 after post-conviction DNA testing established that the hairs in the figure actually originated from different sources.²⁸⁵

TABULAR OR GRAPHIC MATERIAL SET FORTH AT THIS POINT IS NOT DISPLAYABLE

*76 These types of visual demonstrations of microscopic “matches” were supported by impressively credentialed “scientists,” who--like the bite mark experts discussed above--urged jurors to accept their conclusions as probative **evidence** of the defendant's guilt.²⁸⁶ Yet, instead of developing a statistical basis to provide scientifically valid conclusions concerning the probative value of an association between a known and a questioned hair (hypothesis two),²⁸⁷ FBI examiners used meaningless *77 terms of art²⁸⁸ that simultaneously masked the lack of population data while conveying to jurors that the defendant was the likely source of the questioned hair.²⁸⁹ To do so, agents from the FBI's hair and fiber unit proffered--and trained other analysts to proffer²⁹⁰--scientifically false and misleading testimony. This invalid testimony fell into three broad categories. First, and most brazenly, experts baldly asserted and or implied that the defendant was the source of the questioned hair to exclusion of all other potential sources.²⁹¹ Second, experts also used fictitious numbers to assign a statistical weight or probability to the proposition that the questioned hair originated from a particular source.²⁹² Where experts refrained from using numbers, probabilistic opinions were expressed through language that led jurors to believe that valid statistical weight could be assigned to a microscopic hair association.²⁹³ Finally, analysts employed inappropriate substitutions of heuristically gained knowledge for a valid statistical basis in order to bolster the conclusion that a questioned hair belonged to the defendant.²⁹⁴ As in the Odom case discussed above, such testimony was expressed by citing the number of hair analyses the expert performed in the lab over the course of her career and the number of samples from different individuals that could not be *78 distinguished from one another.²⁹⁵ This testimony was proffered to suggest--without **evidence**--that human hair was virtually unique and, therefore, that the microscopic “match” between the known and questioned hair was highly probative **evidence**.²⁹⁶

These misleading statements--accompanied by impressive visual and oral explanations of points of microscopic comparison--allowed experts to argue that human hairs were “microscopically identical,”²⁹⁷ (i.e., essentially unique) and conclude that the questioned hair was therefore “consistent with” having originated from the defendant.²⁹⁸ *79 Affirming appellate courts, however, often pointed to the term “consistent with” in finding that the **evidence**--even if unreliable--was unlikely to have affected the outcome of trial because the expert qualified the conclusions.²⁹⁹

The prosecution of Timothy Scott Bridges in North Carolina is a useful illustration of the dissonance between what a jury was lead to believe the probative value of hair **evidence** was and the way appellate courts characterized the **evidence** on appeal. Such characterizations of hair **evidence** by appellate courts perpetuated decades of invalid precedent and gave license to hair examiners to continue to mislead juries. At Bridges' 1991 trial for the beating and rape of an elderly woman, an FBI-trained hair expert testified that two head hairs found at a crime scene were “consistent with” Bridges.³⁰⁰ He explained that “if an unknown hair is consistent with the standard source in all respects, then it is likely that the hair originated from the same source as the standard.”³⁰¹ While acknowledging that no single characteristic in Bridges' hair could be considered unique, the expert testified that “all of those characteristics in combination is what makes it a strong identification.”³⁰² The analyst invented statistical **evidence** to falsely claim that, based on his examination of

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE..., 4 Va. J. Crim. L. 1

2,000 to 3,000 hairs, the “a conservative estimate ... would be approximately one in a thousand” that one would find two people from the general population with Caucasian head hairs whose hair would contain identical microscopic characteristics.³⁰³ Finally, he testified that it was his opinion that “it is likely that [the two hairs] originated from Timothy Scott Bridges.”³⁰⁴ The appellate court-- relying on the “disclaimer” and the court's previous acceptance of *the same type of* *80 *testimony from the same expert*--found that the statement that “it [the hair] is quite likely to have been from [the defendant],” was appropriate because the expert “did not rule out the possibility that the hair originated from someone other than the defendant.”³⁰⁵ Although the “statistical illustration,” was error, it was harmless because it “was based on the expert's experience and expertise in the hair microscopy field and did not eliminate the possibility of sources of the hair other than defendant.”³⁰⁶ The error was the only physical **evidence** tying Bridges to the crime, and the prosecutor argued that two hairs from the crime scene “match[ed]” the defendant and that the expert--who “was qualified by the judge as an expert in this field of hair”--had assured the jury that “was likely to have come from the defendant.”³⁰⁷

Since Bridges' conviction,³⁰⁸ at least seventy-four defendants have been **wrongfully convicted** based at least in part on microscopic hair comparison.³⁰⁹ While the DNA revolution that inspired Saks and Koehler's predicted paradigm shift has brought some level of additional scrutiny to the misleading claims of these and other trace **evidence** *81 experts,³¹⁰ a century of case law supported--and continues to support--the admissibility of false and misleading expert testimony in the field of hair microscopy.³¹¹ The unanimity of case law proceeded apace, even in jurisdictions with multiple DNA exonerations of **wrongfully convicted** defendants. *State v. Reid*³¹² --the first opportunity the Connecticut Supreme Court had to consider the admissibility of hair comparison **evidence** after the state adopted the *Daubert* test³¹³ --is illustrative of how invalid precedent continues to frustrate due process. In response to a rare and strenuous objection to introduction of hair microscopy **evidence**, the trial court--in still a rarer occurrence--held a *Daubert* hearing prior to the introduction of the **evidence**.³¹⁴ In finding the **evidence** properly admitted because the “technique ha[d] been admitted in Connecticut courts for many years,”³¹⁵ the Connecticut high court went a step further, finding that, in any event, hair comparison **evidence** was not the type of **evidence** required to undergo *Daubert* scrutiny:

Although [the expert's] training is based in science, he testified about a subject that simply required the jurors to use their own powers of observation and comparison. During his testimony, [the expert] displayed an enlarged photograph of one of [Mark Reid's] hairs and one of the hairs recovered from the victim's clothing as they appeared *82 side-by-side under the comparison microscope. [The expert] explained to the jurors how the hairs were similar and what particular features of the hairs were visible. He also drew a diagram of a hair on a courtroom blackboard for the jurors. The jurors were free to make their own determinations as to the weight they would accord the expert's testimony in the light of the photograph and their own powers of observation and comparison. The jurors were not subject to confusing or obscure scientific **evidence**, but were able to use the testimony to guide them in their own determination of the similarity of the two hairs.³¹⁶

Post-conviction DNA testing proved that Reid was innocent of the rape for which he was convicted.³¹⁷ Apart from highly suspect eyewitness identification **evidence**,³¹⁸ the only **evidence** introduced against him was *83 hair comparison testimony from the lead criminologist at the Connecticut Forensic Science Laboratory.³¹⁹ The expert testified “unequivocally” that three pubic hairs discovered on the victim's panties, bra, and sock “were Negroid pubic hairs,”³²⁰ and concluded “‘to a ‘reasonable degree of scientific certainty,’ that the pubic hairs found on the victim's clothing were microscopically similar to those pubic hair samples taken from Mark Reid.”³²¹ DNA testing not only excluded Reid as the source, but it also established that the pubic hairs originated from the caucasian victim.³²² Because

it was a case of first impression, the *Reid* court looked to other jurisdictions for support “that a *Daubert* hearing [was] not required for admission of microscopic hair analysis,”³²³ including the Tenth Circuit's opinion in *Williamson v. Ward*, a decision arising out of a habeas corpus petition by condemned Oklahoma prisoner Ronald Williamson.³²⁴ Like Reid, Williamson was innocent but had been convicted on hair microscopy **evidence**; indeed, Williamson was exonerated the year *before* the *Reid* decision.³²⁵ In Reid's case, the court mischaracterized *Williamson* as standing for proposition that hair microscopy was not the type of ***84 evidence** required to undergo *Daubert* scrutiny.³²⁶ The Tenth Circuit reversed the district court's decision, which found that hair microscopy failed *every element* of the *Daubert* test,³²⁷ because due process--not *Daubert*--was the controlling standard for federal habeas review of state court evidentiary rulings.³²⁸ The Tenth Circuit affirmed the district court's granting of relief on other grounds, but directed the trial court to conduct its own evidentiary hearing on the admissibility of the hair **evidence**.³²⁹

Before the case could be retried, though, Ronald Williamson and Dennis Fritz--a co-defendant who had been tried separately--were exonerated after post-conviction DNA testing revealed that the hair microscopy claims were misleading and false and that a witness for the state had been the true perpetrator.³³⁰ State courts--including Oklahoma criminal courts-- nevertheless took pains to point out that the failed *Daubert* test in *Williamson* was not controlling precedent and thereafter continued to admit hair comparison **evidence**.³³¹ Oklahoma hair ***85** microscopy jurisprudence demonstrates the lethal dissonance between scientific reality and legal precedent. The rejection of the federal court's *Daubert* analysis occurred in 1997;³³² between 1998 and 2012, eight defendants whose convictions were obtained at least in part through hair **evidence** were exonerated, including three people sentenced to death.³³³ Nonetheless, there has never been a negative admissibility ruling concerning such hair **evidence** in Oklahoma; in the only documented challenge, the defendant lost.³³⁴ The defendant in that case--Curtis McCarty--was innocent of the capital murder for which he was convicted.³³⁵ He argued for the exclusion of hair **evidence** in a pretrial motion *in limine* because “the reliability of forensic hair comparison ***86 evidence** has not been adequately established.”³³⁶ The court admitted the **evidence**, and the appellate court, relying on Oklahoma precedent, reaffirmed the admissibility of hair comparison **evidence**.³³⁷

Perhaps most interesting is courts' treatment of the admissibility of hair **evidence** in those jurisdictions that featured precedent *excluding* the discipline. Such treatment is, in some sense, a disturbing inversion of trial and appellate courts' reflexive embrace of precedent in the bite mark context. In New Hampshire, for example, the state supreme court affirmed the trial court's exclusion of the **evidence** as early as 1969, noting that “the **evidence** on hair-identification offered by the State would not be acceptable to scientists in the field” and, as such, did not meet the *Frye* standard requirement that the “scientific principle involved ... ‘be sufficiently established to have gained general acceptance in the particular field in which it belongs.’”³³⁸ Thereafter, the **evidence** was admitted in three reported New Hampshire cases decided subsequent to *Coolidge*--in 1976, 1978 and 1981 respectively--which dealt with the admissibility of hair testimony.³³⁹ In *People v. Roff*, a 1979 New York case, the appellate court found that:

***87** [The trial court erred] in receiving and refusing to strike the testimony of the chemist that the hair taken from the bathroom and found at the scene of the crime could have come from the same person and that there was some similarity between the two samples, and erred in receiving the physical **evidence** itself ...

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE..., 4 Va. J. Crim. L. 1

because the **evidence** was inadequate to connect the hair samples with defendant's hair, it was inadmissible, because it did not “accurately portray a relevant and material element of the case.”³⁴⁰

Thereafter, however, New York hair **evidence** jurisprudence stands in direct opposition to *Roff*, supporting the admissibility of such **evidence**.³⁴¹

C. THE FBI'S TRAINING OF HUNDREDS OF STATE HAIR EXAMINERS TO PROVIDE FALSE AND MISLEADING TESTIMONY

The false hair comparison **evidence** that helped **wrongfully convict** David Johns Bryson and Jimmy Ray Bromgard was introduced through two disgraced forensic analysts from state crime labs: Joyce Gilchrist and Arnold Melnikoff, respectively.³⁴² Both analysts--like many hundreds of others, including the analyst in the Bridges case³⁴³--were trained by the FBI in a two-week training course at the FBI Academy in Quantico, Virginia, and both later stated that they had been trained by the FBI to *88 provide false and misleading testimony.³⁴⁴ Gilchrist's records showed that she was trained to use her experience to bolster the strength of her conclusion:

In her files, she kept a certificate of completion from her January 1981 class, including a session on ‘Discussion of the significance of hair comparisons, testimony matters and pertinent literature.’

In her notes, she copied the FBI caveat that one cannot conclusively determine the source or origin of a hair. But, the notes also showed that instructors were teaching their students how to sidestep the limits of the science--by pointing out their experience.

“Can conclude source--point out however in my experience, have rarely seen hairs from diff. people exhibiting the same microscopic characteristics,” the notes say.³⁴⁵

Other FBI-trained examiners made nearly identical assertions related to using bench experience as both a workaround for the lack of a valid statistical basis and a way to undermine the “disclaimer.”³⁴⁶ As noted, there is **evidence** that there was, in fact, no massive failure by FBI scientists to appreciate fundamental scientific principles, but rather that the limitations of hair comparison **evidence** were understood and deliberately obscured to implicate defendants. For example, during the 1985 “International Symposium on Forensic Hair Comparisons” at the FBI Academy in Quantico, a revealing panel discussion took place concerning the lack of a statistical basis to support the claims of hair comparison *89 experts.³⁴⁷ This panel included two participants from the FBI Laboratory, the Chief Scientist for the Hair and Fibers Section for the Royal Canadian Mounted Police, the Chief Scientific Officer for the Metropolitan Police Forensic Science Laboratory in London, the Scientific Director for the General Biology Section in Germany, and Dr. Peter De Forest, a Professor of Criminalistics at John Jay College.³⁴⁸ During the discussion, Dr. De Forest explained the limits of the “**evidential** value of hair” and some of the “defense expert's perspectives on the hair question” that he had gleaned through his experience.³⁴⁹ He emphasized that “hair examination” is non-absolute associative **evidence** whose “power”

is to “exclude hair.”³⁵⁰ Dr. De Forest also explained how experts are prone to overstating the value of hair comparison through the inappropriate leveraging of bench experience:

I have a problem with the divergence from a laboratory report in which the conclusion is these hairs could have shared a common origin to the presentation of testimony in court when the expert says something to the effect that, “Yes, these hairs were found to be similar and in my experience I have examined thousands of hairs and I have never found two hairs from different sources that were alike.” I think that is very misleading and it is not substantiated by any data.³⁵¹

Dr. De Forest and other panelists emphasized that it was “clear more [had] to be done” concerning the training of hair microscopy *90 “experts.”³⁵² He noted that he was involved with an FBI-sponsored Committee of Forensic Hair Comparison, which he felt “should be an ongoing committee” because they had “not solved all the problems by any means.”³⁵³ More recently, FBI Special Agent Michael Malone--a former supervisor in the FBI's Hair and Fiber Unit--acknowledged in a civil deposition that, by the mid-1980s, FBI agents had had conversations “to [the] effect” that “[s]ince we didn't have a database and we didn't have, you know, real probabilities, scientifically valid probabilities, let's try and use these numbers of the cases that we have looked at in lieu of real probabilities.”³⁵⁴

Nevertheless, as recently as 2009, the FBI published a report insisting that the technique could provide a “strong basis for an association” and that, although hair microscopy was “not a means of positive identification,” it could “provide substantial information because of the variation in hair among individuals.”³⁵⁵ The report likewise stated that the significance of the association could be expressed “qualitatively or semiquantitatively” and that the only “limitation on the science [was] that there [was] always the possibility of a coincidental match.”³⁵⁶

*91 D. THE FALL OF HAIR MICROSCOPY

The tide only began to turn against hair microscopy with the 2009 publication of the NAS Report.³⁵⁷ The Report concluded that hair microscopy could not uniquely identify one person as the source of a hair; instead, at best, a “match” between two hair samples “mean[t] only that the hair *could* have come from any person whose hair exhibited - within some levels of measurement uncertainties - the same microscopic characteristics.”³⁵⁸ In addition, the NAS Report made it clear that first hypothesis underlying the technique was invalid, noting that there was no consensus “on the number of features on which hairs must agree before an examiner may declare a ‘match.’”³⁵⁹ The second hypothesis was likewise found to be invalid, as there were no statistics on the distribution of particular hair characteristics in the population.³⁶⁰ The Report's ultimate conclusion was “that testimony linking microscopic hair analysis with particular defendants [was] highly unreliable,” and that **evidence** of a match “must be confirmed using mtDNA analysis.”³⁶¹

Following the NAS Report, a series of news articles in *The Washington Post* revealed that erroneous testimony by FBI hair examiners was “widespread and could affect potentially thousands of cases in federal, state and local courts.”³⁶² *The Post* reported that, despite the claims made by FBI analysts, it was virtually impossible to distinguish between two human hairs from different sources; in one instance, mitochondrial DNA testing revealed that two “FBI-trained analysts ... could not even distinguish human hairs from canine hairs.”³⁶³ Consequently, “hundreds of defendants nationwide remain in

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE..., 4 Va. J. Crim. L. 1

prison or on *92 parole for crimes that might merit exoneration, a retrial, or a retesting of **evidence** using DNA because FBI hair and fiber experts may have misidentified them as suspects.”³⁶⁴

The NAS Report, in conjunction with intense media scrutiny and numerous exonerations--particularly those of three men in the Washington, D.C. area who were exonerated after mitochondrial DNA contradicted hair microscopy testimony proffered by FBI examiners³⁶⁵--eventually persuaded the FBI to reexamine the thousands of cases between 1980 and 2000 where its agents testified to a positive association between a defendant's hair and a questioned hair collected from a crime scene.³⁶⁶ In doing so, the FBI essentially adopted the NAS Report's critique, acknowledging for the first time that hair microscopy is limited “in that the size of the pool of people who could be included as a possible source of a specific hair is unknown.”³⁶⁷ Therefore, an examiner may not apply “probabilities to a particular inclusion of someone as a source of a hair of unknown origin.”³⁶⁸ Instead, the strongest conclusion one can draw is that the suspect could be included in a class of people of unknown size that *93 could have been a possible source of the evidentiary hair.³⁶⁹

The FBI has acknowledged the three categories of errors discussed above,³⁷⁰ which were routinely proffered by its agents--and those they trained--in thousands of cases.³⁷¹ In the FBI's ongoing review, the DOJ has agreed to waive any statute of limitations or procedural bars under 28 U.S.C. § 2255 in federal post-conviction cases where error is found, due to the manifest unfairness of punishing a criminal defendant for the FBI's protracted campaign to advance the prosecution's case without regard to the limitations of hair microscopy.³⁷² Even so, hair comparisons have been proffered to juries as “scientific” **evidence** and used to convict people--and to uphold **wrongful convictions** on direct appeal--sometimes with very little other corroborating or incriminating **evidence**.

IV. ATTENDANT OBLIGATIONS

The long-predicted and now manifest paradigm shift in forensic identification **evidence** is rooted in a systemic, century-long failure by nearly all criminal justice stakeholders to comprehend, question, challenge, and exclude as unreliable the false and misleading assertions made by forensic experts and exploited by advocates to persuade lay jurors. This fundamental breakdown in the adversarial process this nation relies upon to discover the truth--illustrated most starkly by the empirical data--compels several prospective and retrospective ethical and professional obligations. Because the shift has occurred across disciplines and over time, the obligations extend not only to specific individuals in those disciplines, but also to distinct professional and governmental entities.

What follows in this section are some broad suggestions for how our justice system might recover scientific integrity and how these solutions might be implemented. Included as well as are some discrete *94 proposals that we believe must be an aspect of whatever solution is implemented. Some of what is at stake is obvious, beginning with the reputation of the criminal justice system itself.³⁷³ If past history is any indication, it goes almost without mentioning that a failure to act will stymie the discovery of numerous instances of **wrongful conviction**. Apart from that, though, as former Attorney General Janet Reno once noted at a conference to address the **wrongful conviction** phenomenon:

If the public's confidence in the results of the criminal justice system erode, then the public will not accept the criminal justice system's findings and results ... [a]nd what we do with the criminal justice system, which is the hallmark of the legal system for so many Americans looking in from the outside, will make a profound difference for this century.³⁷⁴

The ethical and professional obligations that we argue apply are unprecedented. But so too are the circumstances that obligate their imposition. The system-wide problems that we have identified here will not self-correct.³⁷⁵ Additionally, traditional sources of authority for corrective direction--including the Model Rules and governing bodies' ethical guidelines-- are almost exclusively prospectively focused and, more importantly, simply do not address these kinds of systemic failures. While there have certainly been instances where courts have attempted to correct episodes of systemic forensic fraud, those instances were less about the shortcomings of a discipline and more about the malfeasance of *95 specific individuals involved in them.³⁷⁶ Here, the fault lies at the core of the disciplines themselves and in layers of invalid legal precedent. There is no rule or ethical obligation that contemplates these problems, much less addresses them. To the extent that individuals are involved, it is not discrete outliers, but line prosecutors and defense attorneys, who acted without correction for decades. Those prosecutors and defense attorneys were aided and abetted by forensic witnesses--who operated within unvalidated disciplines, exaggerated findings, or both--as well as courts that shamelessly facilitated all of it, abdicating their gatekeeping responsibilities and relying on lay jurors to separate science from subjective speculation convincingly masquerading as science. At the same, courts continued to allow courtroom advocates to further exploit invalid expert opinions in opening and closing statements.³⁷⁷ Where DNA exonerations made it apparent that the claims these experts were making were grossly unreliable, courts failed to engage in any type of responsible analysis when asked to provide correction. The list of those held accountable for proffering unvalidated forensic testimony is short;³⁷⁸ for *96 those individuals and entities responsible for the admission of bite mark testimony and hair microscopy **evidence**, the list is virtually non-existent.

For these reasons, then, we first argue that at least two positions typically proffered by prosecutors to deny petitioners post-conviction relief ought to be unethical to advance as a result of the information marshaled in this article. It is our position that individual prosecutors should be ethically barred from arguing: (1) that a defendant-petitioner should have known at the time of trial that a forensic discipline was unvalidated or false, particularly if its admission into **evidence** was achieved through improper processes, like the ones described earlier; and, (2) that whatever prejudice may have resulted from the admission of such **evidence** could have been cured through cross-examination and/or by the "disclaimers" described above.³⁷⁹ Second, as it relates to defense counsel, we argue that it is per se ineffective assistance of counsel³⁸⁰ to fail to *98 challenge these unvalidated disciplines going forward. Third, as it relates to courts, we argue that: (1) taking judicial notice of the admissibility of putative scientific **evidence** is inappropriate because science is not static,³⁸¹ (2) decisions regarding the admissibility of trace **evidence** should be treated as cases of first impression, without any reliance on flawed precedent; and, (3) similarly, following the DOJ's lead in waiving procedural objections, reviewing courts should not invoke procedural bars to deny relief to defendant-petitioners if the rationale supporting denial of *99 relief is that defendant-petitioners knew or should have known of the disciplines' flaws. Lastly--again following the FBI/DOJ's lead--we argue that there now exists affirmative ethical and professional obligations³⁸² on a host of entities both to identify and review case files for convictions based in whole or in part on unvalidated forensic science and to make substantive contact with affected defendant-petitioners, as well as the final prosecuting body, defense counsel of record, and the tribunals where jurisdiction lies for those cases.

A. UNIQUE NATURE OF ETHICAL & PROFESSIONAL OBLIGATIONS

In contemplating how best to implement these professional and ethical obligations, it is worth discussing briefly why currently available remedies are inadequate. The scholarship on the ethical implications surrounding questionable forensic **evidence** is prolific,³⁸³ especially as it *100 concerns prosecutors' duties.³⁸⁴ There is even specific scholarship and guidance directed at the use of hair microscopy **evidence**³⁸⁵ and bite mark **evidence**.³⁸⁶ But these discussions do

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE..., 4 Va. J. Crim. L. 1

not address the problems that we have illustrated for at least two significant reasons. First, there is a temporal problem. Model Rule of Professional Responsibility 3.3, “Candor Toward the Tribunal,” requires a lawyer not to “offer **evidence** that the lawyer knows to be false”³⁸⁷ and states that, “[i]f a lawyer ... comes to know of ... [the **evidence's**] falsity, the lawyer shall take reasonable remedial measures, including, if necessary, disclosure to the tribunal.”³⁸⁸ But the obligation to take remedial measures extends only *101 until “the conclusion of the proceeding,”³⁸⁹ which the comment to the rule defines as “when a final judgment in the proceeding has been affirmed on appeal or the time for review has passed.”³⁹⁰ In almost every instance, the set of affected cases that is of concern here will fall far outside of the timeframe that would require a lawyer to take remedial measures.³⁹¹ Similarly, other discussions--like those involving amending Rule 3.8,³⁹² *Special Responsibilities of a Prosecutor*, to add a “gatekeeping role” for prosecutors or those invoking Rule 1.1's requirement of “competence”³⁹³ to counter claims that Rule 3.3's, *Candor Toward the Tribunal*, “knowing” scienter and the elastic definition of what comprises “false” **evidence** for advocates prohibits the imposition of ethical sanctions-- are likewise unhelpful. To begin with they are prospective solutions, and, as solutions, seem unlikely either to be pursued or seriously adjudicated.³⁹⁴ More specifically, those responsible for these failures of justice-- or, maybe more importantly, those who would be most effective at addressing them--have not acted based on what are typically viewed as incentives to do so. For example, the doctrine of prosecutorial immunity--which had been at least somewhat limited so that aggrieved petitioners might be able to seek redress for the most abusive acts of prosecutorial malfeasance³⁹⁵ -- has not made a difference in redressing these failures in any meaningful or *102 measurable way. In fact, the Supreme Court has substantially broadened, not limited, the protection afforded to prosecutors by the doctrine.³⁹⁶

State bar discipline--also held out as an incentive--is likewise an unsatisfactory solution.³⁹⁷ Available data is replete with the systemic failure of state bar disciplinary entities to hold prosecutors--or defense attorneys, for that matter--accountable for misconduct.³⁹⁸ Judges, too-- *103 who are arguably best-suited to observe and consider misconduct in the trials over which they preside--fare no better. In a 2008 study in California, data showed that, in cases involving findings of prosecutorial misconduct between 1997 to 2006--of which there were 444--fifty-four were reversed, which triggered a per se legal duty to report.³⁹⁹ Yet there was not a single referral.⁴⁰⁰

Also unlikely to be helpful, at least in the near future, are forensic science organizations themselves. They have shown little to no inclination to address seriously the problems for which they are directly responsible, particularly issues related to fundamental scientific weaknesses. In addition to the sweeping critique of the substance of many traditional forensic science disciplines, the NAS Report also noted that only “some fields” have “[s]tandards and codes of ethics.”⁴⁰¹ Those that do have codes of ethics have codes that “vary in content”⁴⁰² and lack “consistent mechanisms for enforcing”⁴⁰³ them. Furthermore:

Many jurisdictions do not require certification in the same way that, for example, states require lawyers to be licensed. Therefore, few forensic science practitioners face the threat of official sanctions or loss of certification for serious ethical violations. And it is unclear whether and to what extent forensic science practitioners are required to adhere to ethics standards as a condition of employment.⁴⁰⁴

The NAS's concern is borne out by anecdotal **evidence**. As Spencer Hsu of *The Washington Post* reported with respect to forensic hair analysis, for example, even though “Justice Department officials ... [knew] for years that flawed forensic work might have led to the convictions of potentially *104 innocent people ... prosecutors failed to notify defendants or their attorneys even in many cases they knew were troubled.”⁴⁰⁵ According to a July 2014 report by the DOJ Office of the Inspector General, there were several “serious deficiencies” with an FBI Criminal Division Task Force's internal

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE..., 4 Va. J. Crim. L. 1

review of “cases involving the use of scientifically unsupportable analysis and overstated testimony by FBI Lab examiners in criminal prosecutions.”⁴⁰⁶ First among the deficiencies was the failure to prioritize capital cases in its review.⁴⁰⁷ It took the Task Force almost five years to identify affected death penalty cases, thus depriving “state authorities” of the basis “to consider delaying scheduled executions.”⁴⁰⁸ As a result, Texas executed Benjamin H. Boyle⁴⁰⁹ before his case was reviewed by the Task Force even though “[t]he prosecutor deemed the [FBI] Lab analysis and testimony ... material to the defendant's conviction”⁴¹⁰ and death sentence.⁴¹¹ In addition, the OIG report found that the affected defendants were not provided with “appropriate and timely disclosures ... *105 particularly in cases ... [where] the analysis or testimony was material to the conviction and the report of the independent scientists established that such evidence was unreliable.”⁴¹²

The response to the conclusions of the NAS Report and the ever-increasing number of wrongful convictions from the insular and largely independent forensic odontology community demonstrates the urgent need for legislation to provide avenues of post-conviction relief for prisoners whose convictions rest on discredited scientific evidence. First, there has been no effort at all to address known problems in past cases, even though the empirical data that would support such a review is well-documented and mounts annually.⁴¹³ Worse, efforts that have been made to rectify the discipline's shortcomings have been directed mainly at salvaging an increasingly maligned discipline. For example, in August 2013--in the wake of wrongful convictions and indictments, lawsuits against the dentists who proffered false and misleading testimony,⁴¹⁴ and the devastating conclusions of the NAS Report--the American Board of Forensic Odontology (ABFO) finally conceded that individualization claims were invalid in “open population” cases where the universe of potential suspects was unknown.⁴¹⁵ This dramatic and unprecedented *106 change in the guidelines is a long-overdue admission that such testimony is scientifically invalid. But the change was not made publically, and no effort was made by the ABFO, or any other entity, to identify those convictions that were, in whole or in part, the result of this type of now-rejected methodology.⁴¹⁶ In short, the development, such as it is, seems to affect only the argument for the continued legitimacy of the discipline itself.⁴¹⁷

*107 Even assuming, though, a best-case scenario in which actors act with appropriate humility and haste, there are a host of practical difficulties. Hair microscopy serves as a good example. A legitimate audit of cases involving unvalidated hair microscopy evidence would have to include not only cases in which FBI analysts testified--several thousand cases over a twenty-five year period--but also those in which state analysts testified. As discussed, beginning in the late 1970's, the FBI lab implemented a two-week training program in hair and fiber analysis for state and local lab employees, and there is ample evidence state practitioners were taught to proffer misleading testimony to triers of fact.⁴¹⁸ As more and more states began to rely on their local and state labs to provide hair examiner reports and testimony in their state investigations and prosecutions--particularly in the 1980s and 1990s--the Bureau's two-week program trained in excess of 500 examiners over a period of twenty- *108 five years.⁴¹⁹ In short, there are likely thousands of cases--in which FBI examiners or FBI-trained state analysts provided testimony⁴²⁰--that would need to be part of an audit.⁴²¹ Moreover, to the extent that errors are found, counsel and defendants in those cases must be appropriately notified. Many of the cases will be decades old, records will be difficult to locate, and, in some instances, counsel will be impossible to locate. For jurisdictions already strapped for resources to fund their criminal justice systems, finding the resources--monetary and otherwise--to do so may well turn out to be impossible.

*109 B. SUGGESTED SOLUTIONS

Conviction Integrity Programs (CIPs)⁴²² have been used with success throughout a number of prosecutors' offices around the country, and with certain modifications, may be the best administrative template for a successful remedy. Foremost among these programs' strengths is their practical approach to reviewing a discrete, identifiable set of cases for specific types of error. In essence, CIPs fill a gap that Rule 3.8 does not.⁴²³ In addition, they can stand as real, on-the-ground embodiments of aspirational standards that exist elsewhere.⁴²⁴ The units that we propose would-- given the discrete focus on a certain subset of cases involving unreliable scientific **evidence**--be developed outside of, rather than within, prosecutors' offices and would thus function more like a neutral administrative agency than a branch of an adversarial office.⁴²⁵ Like other successful CIPs, their founding structure would incorporate best practices that, among other things, would grant them privileged access and cooperation, namely open file sharing-- including work product from both prosecutors' and defense attorneys' files--and mutual investigative assistance from all individuals and entities, including the forensic labs and analysts⁴²⁶ involved in identified cases.

*110 C. MODIFICATION OF PROCEDURAL BARRIERS

In addition to these modified conviction integrity models, procedural barriers cannot be erected to frustrate the very purpose of auditing these cases: the determination of whether false scientific **evidence** contributed to securing a conviction. Thus, where such **evidence** was introduced, waivers of typical statute of limitations bars and other procedural default mechanisms must be granted as a matter of course. These suggestions are in accord with newly-developed post-conviction statutory modifications adopted in Texas and California, and, equally as importantly, coincide with the position that the DOJ has adopted with respect to affected cases identified in its hair microscopy audit.⁴²⁷ With respect to the FBI/DOJ audit, for example, letters notifying parties of the introduction of false **evidence** have stated:

In the event that the defendant seeks post-conviction relief based on the Department's disclosure that microscopic hair comparison reports or testimony used in this case contained statements that exceeded the limits of science, we provide the following information to make you aware of how we are handling such situations in federal cases. In such cases under 28 U.S.C. § 2255, in the interest of justice, the United States is waiving reliance on the statute of limitations under Section 2255(f) and any procedural default defense in order to permit the resolution of legal claims arising from the erroneous presentation of microscopic hair examination laboratory reports or testimony.⁴²⁸

*111 In addition, most states' post-conviction statutes require not simply that **evidence**--in this case, errors affecting the admitted forensic **evidence**--is newly-discovered, but that its discovery would have affected the outcome of the trial.⁴²⁹ This standard, too, should be modified with respect to the cases that a CIP deems meritorious. Specifically, because so many of these cases will be so old and information difficult to access and assess, the standard should be akin to a due process analysis of whether false **evidence** was admitted into the trial, and, if so, whether there is any reasonable likelihood the **evidence** affected the judgment of the jury.⁴³⁰ If so, relief should be warranted. Alternatively, the burden of proving that the trial was fundamentally fair notwithstanding the introduction of unvalidated forensic **evidence** should rest with the prosecution, which would be required to show that the constitutional error was harmless "beyond a reasonable doubt." More specifically, where the court, the prosecutor, and defense counsel all operated under the false assumption

*112 that the scientific **evidence** at issue was valid and reliable, there was no meaningful adversarial testing of the

false **evidence**. Thus, the introduction of the now discredited **evidence**--which was nevertheless proffered to the jury as infallible "scientific" **evidence** of guilt--was so unfair it resulted in a "breakdown in the adversarial process" in violation of petitioner's due process rights.⁴³¹

CONCLUSION

On one hand, the extent of the problems that this article illustrates and the call that it makes for affirmative acts of reform risk its being characterized as simply more of the same: a partisan philosophical position about the state of the criminal justice system, albeit this time costumed with an abundance of data and excerpts from case law. Were that characterization correct, then equally partisan responses in opposition could follow as a matter of course. The end result would be a stalemate: *113 one side arguing that what this article documents is the natural by-product of a broken system; the other that it is the natural, collateral consequence of a system trying--albeit with too much aspiration--to balance public safety against the competing claims of defendants' due process rights.

The fact of the matter is, however, that this article, though it certainly documents disturbing failures--both in individual cases as well in several disciplines--is nevertheless focused on a finite number of specific cases, a circumscribed jurisprudence, and a group of individuals and entities that can themselves provide an immediate and effective solution. Or not. The results of that decision, though, are stark. To the extent that it is an overstatement to claim that a decision one way or the other defines the character of the system as a whole, it is not too much to claim that--given what we know about the kinds of failures documented here--a decision to act, or not, characterizes specific individuals and entities. And that characterization works from the bottom up, as it were, to create a larger, more resonant definition.

To illustrate, consider the following case: The defendant was convicted of sexual assault in Mississippi in 1981 and sentenced to twenty-five years in prison. The **evidence** against him, as the Mississippi Supreme Court noted, "was conflicting."⁴³² He was identified by the victim as the person who had assaulted her, as well as by another individual, who testified that she had observed the defendant at the victim's house on the day of the assault.⁴³³ The defendant denied having committed the offense and testified that he had been in Chicago on the day it occurred.⁴³⁴ He voluntarily surrendered to police upon his return.⁴³⁵

The only physical **evidence** that connected the defendant to the crime scene was human hair.⁴³⁶ According to the court, "[h]air samples taken from the appellant and the prosecutrix's clothing were compared in *114 the F.B.I. laboratory. All twenty individual characteristics identified in appellant's hair matched the characteristics of the hair taken from the victim's clothing."⁴³⁷

In 2001, after the DOJ and FBI became aware that the analyst who had provided the testimony in the Mississippi case was Michael Malone--known by then "as the agent making the most frequent exaggerated testimony"⁴³⁸--the DOJ wrote a letter to the district attorney in Mississippi whose office had prosecuted the case.⁴³⁹ The letter alerted the prosecutor to the fact that the Mississippi case was under federal review and asked the prosecutor for "any other information [he] may have related to the ... case to determine if Malone's laboratory work was material to the conviction."⁴⁴⁰

By that time, the case had been appealed and affirmed, and the trial transcript--at least the copy that the Mississippi Supreme Court had used--was located in the State archives in Jackson. The authors recently read it. Among the claims that the FBI analyst Malone made were these: that in order to be qualified for his job, he had to perfectly match fifty hairs to fifty people⁴⁴¹ and that the hairs recovered from the crime scene "microscopically matched the head hairs of ... [the

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE..., 4 Va. J. Crim. L. 1

defendant]. In other words, they were indistinguishable from his head hairs. How unlikely [would it be for two different people to share the same observed characteristics]? In about ten thousand hair exams, I've only seen two occasions where I had hairs from two different people that I couldn't distinguish.”⁴⁴²

In March of 2002--eight months after the DOJ had alerted the *115 district attorney to the potential problem and asked for assistance--the district attorney responded. In a handwritten response on a single fax cover page, the district attorney said “This is a 20 year old case with all record files having been previously destroyed. No determination to your request can be made.”⁴⁴³

No substantive additional action has been taken on the case since.

Footnotes

- 1 Joseph Flom Special Counsel & Director of Strategic Ligation for the Innocence Project.
- 2 Professor and Director of the Mississippi Innocence Project and Clinic at the University of Mississippi School of Law. This article benefited greatly from the contributions of Sarah Krieger, Amelia Maxfield, Joanne Luckey, and Marielle Dirx. The authors owe a debt of gratitude to the staff of the Virginia Journal of Criminal Law for their heroic editorial work on this article, particularly Amelia Nemitz. Finally, many thanks to Professor Brandon Garrett for guidance on this project and for his transformative scholarship on the intersection of science, law, and wrongful conviction.
- 3 Jonathan Koehler & Michael Saks, *The Coming Paradigm Shift in Forensic Identification Science*, 309 SCIENCE 892, 892 (2005).
- 4 See NAT'L RESEARCH COUNCIL OF THE NAT'L ACADEMIES OF SCIENCES, STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD (2009), available at <http://www.nap.edu/catalog/12589.html> [hereinafter NAS REPORT].
- 5 *Han Tak Lee v. Tennis*, No. 4:08-CV-1972, slip op. at 1, 1-2 (M.D. Pa. June 13, 2014) [hereinafter *Han Tak Lee II*]. As the court explained, “[t]his proverb, inscribed at the University of Pennsylvania Law School on the statue of Hseih-Chai, a mythological Chinese beast who was endowed with the faculty of discerning the guilty, is a fitting metaphor for both the progress of the law and the history of this case.” *Id.* at 1.
- 6 *Id.* at 2.
- 7 Michael Rubinkam, *Arson Science Discredited*, COLUMBUS DISPATCH (June 23, 2014), http://www.dispatch.com/content/stories/national_world/2014/06/23/arson-science-discredited.html.
- 8 *Han Tak Lee II*, No. 4:08-CV-1972, slip op. at 1.
- 9 See generally Rachel Dioso-Villa, *Scientific and Legal Developments in Arson Investigation Expertise in Texas v. Willingham*, 14 MINN. J.L. SCI. & TECH. 817 (2013); David Grann, *Trial by Fire: Did Texas Execute an Innocent Man?*, NEW YORKER (Sept. 7, 2009), <http://www.newyorker.com/magazine/2009/09/07/trial-by-fire>.
- 10 See generally Brandon L. Garrett & Peter J. Neufeld, *Invalid Forensic Science Testimony and Wrongful Convictions*, 95 VA. L. REV. 1 (2009); Caitlin Plummer & Imran Syed, “*Shifted Science*” and *Post-Conviction Relief*, 8 STAN. J. C.R. & C.L. 259 (2012).
- 11 *Han Tak Lee II*, No. 4:08-CV-1972, slip op. at 1.
- 12 Koehler & Saks, *The Coming Paradigm Shift in Forensic Identification Science*, *supra* note 3.

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE..., 4 Va. J. Crim. L. 1

- 13 KELLY M. PYREK, *FORENSIC SCIENCE UNDER SIEGE: THE CHALLENGES OF FORENSIC LABORATORIES AND THE MEDICO-LEGAL INVESTIGATION SYSTEM* 222 (Elsevier Acad. Press 2007).
- 14 [State v. Stinson](#), 397 N.W.2d 136, 138 (Wis. Ct. App. 1986).
- 15 Transcript of Record at 83, [State v. Stinson](#), 397 N.W.2d 136 (Wis. Ct. App. 1984) (No. 86-0002-CR).
- 16 [Stinson](#), 397 N.W.2d at 138.
- 17 *Id.* Stinson was exonerated in 2009. See [Robert Lee Stinson](#), INNOCENCE PROJECT, http://www.innocenceproject.org/Content/Robert_Lee_Stinson.php (last visited Oct. 9, 2015).
- 18 Transcript of Record at 730, [State v. Brooks](#), 748 So. 2d 736 (Miss. 1992) (No. 98-KA-00322-SCT).
- 19 See [Robert Lee Stinson](#), *supra* note 17.
- 20 See [Levon Brooks](#), INNOCENCE PROJECT, <http://www.innocenceproject.org/cases-false-imprisonment/levon-brooks> (last visited Oct. 9, 2015).
- 21 The governing standard for expert opinion admissibility in Stinson's case was that “evidence given by a qualified expert is admissible irrespective of the underlying scientific theory ... [as long as the] expert scientific testimony [is] relevant ... and the expert [is] qualified [so that] scientific or specialized knowledge will assist the trier of fact to determine a fact in issue” [Stinson](#), 397 N.W.2d at 140-41 (citing [State v. Walstad](#), 351 N.W.2d 469 (Wis. 1984)). In [Brooks v. State](#), the Mississippi Supreme Court simply adopted a blanket rule in favor of the admission of bite mark testimony, but noted that “[i]t is certainly open to defense counsel to attack the qualifications of the expert, the methods and data used to compare the bite marks to persons other than the defendant, and the factual and logical bases of the expert's opinions. Also, where such expert testimony is allowed by the trial court, it should be open to the defendant to present evidence challenging the reliability of bite-mark comparisons.” [Brooks v. State](#), 748 So. 2d 736, 739 (Miss. 1999) (quoting [Howard v. State](#), 701 So. 2d 274, 288 (Miss. 1997) (No. 94-DP-00524-SCT)) (emphasis in original).
- 22 [Brooks](#), 748 So. 2d at 746 (Smith, J., concurring).
- 23 *Id.* at 739.
- 24 In a companion case, Kennedy Brewer was exonerated a few weeks before Brooks. See [Levon Brooks](#), *supra* note 20. Bite mark evidence testimony from the same state expert had also led to Brewer's conviction, and, like Brooks' exoneration, was debunked as a result. See *id.*; Coburn Dukehart, *Flawed Autopsies Send Two Innocent Men to Jail*, NPR, <http://www.npr.org/2011/06/01/133401716/flawed-autopsies-send-two-innocent-men-to-jail> (last visited Feb. 13, 2016). Post-conviction DNA testing identified the true perpetrator, who had murdered the victims in each case. See [Kennedy Brewer](#), INNOCENCE PROJECT, <http://www.innocenceproject.org/cases-falseimprisonment/kennedy-brewer> (last visited Oct. 9, 2015); [Levon Brooks](#), *supra* note 20.
- 25 See *infra* note 145 and accompanying text.
- 26 See, e.g., [Han Tak Lee v. Tennis](#), No. 4:08-CV-1972, slip op. at 1, 3-5 (M.D. Pa. Sept. 22, 2010) [hereinafter *Han Tak Lee I*]; [Willingham v. Johnson](#), No. 02-10133, 2001 WL 1677023, at *1 (N.D. Tex. Dec. 31, 2001); [Commonwealth v. Lee](#), 911 A.2d 182 (Pa. Super. Ct. 2006) (No. 3237 EDA 2005).
- 27 See *Han Tak Lee I*, No. 4:08-CV-1972, slip op. at 1.
- 28 28 U.S.C.A. § 2254 (West 1996).
- 29 *Han Tak Lee I*, No. 4:08-CV-1972, slip op. at 5. The court cited [Herrera v. Collins](#), 506 U.S. 390, 400 (1993) for the following proposition:

Claims of actual innocence based on newly discovered **evidence** have never been held to state a ground for federal habeas relief absent an independent constitutional violation occurring in the underlying state criminal proceeding “This rule is grounded in the principle that federal habeas courts sit to ensure that individuals are not imprisoned in violation of the Constitution—not to correct errors of fact.”

See *Han Tak Lee II*, No. 4:08-CV-1972, slip op. at 5 (quoting *Herrera*, 506 U.S. 390, 400 (1993)).

30 Unfortunately, this claim may be overstated as it relates to *Willingham*. See, e.g., Maurice Possley, *Fresh Doubts over a Texas Execution: New **Evidence** Revives Concerns That a Man Was Wrongly Put to Death in 2004*, WASH. POST (Aug. 3, 2014), <http://www.washingtonpost.com/sf/national/2014/08/03/fresh-doubts-over-a-texas-execution/?hpid=z1>.

31 With respect to death penalty cases alone, we have identified at least fifteen convictions where bite mark **evidence** not only played a key role in the prosecution, but where, as in *Stinson* and *Brooks*, the cases also mutually relied on each others' flawed acceptance of the pseudo-science used to justify the conviction. See generally INNOCENCE PROJECT, STRATEGIC LITIGATION UNIT, DEATH PENALTY CONVICTIONS SUPPORTED BY BITE MARK **EVIDENCE** (2014) (on file with authors).

32 853 So. 2d 781 (Miss. 2003).

33 *Id.* at 784-85.

34 The Mississippi Supreme Court reversed Mr. Howard's first conviction and remanded the case for a new trial, finding that Mr. Howard should not have been allowed to proceed *pro se* in light of his apparent mental condition and because it was a capital case. See *Howard v. State*, 701 So. 2d 274, 274, 283, 288 (Miss. 1997).

35 See *Howard*, 853 So. 2d at 795-96; *id.* at 799-800 (McRae, J., dissenting).

36 See *id.* at 795-96.

37 Response in Opposition to Petitioner's Motion to Vacate Conviction, and Supplemental Memorandum in Support of Post-Conviction Relief at 45, *Howard v. State*, No. 2014-DR-01305-SCT (Miss. Jan. 14, 2015).

38 *Id.*

39 *Howard v. State*, 945 So. 2d 326, 363 (Miss. 2006) (No. 2003-DR-01881-SCT).

40 The Innocence Project's Strategic Litigation Unit uses litigation to challenge judicial reliance on **evidence** from unreliable forensic science disciplines and to reform the legal framework used to evaluate eyewitness identification **evidence**. See *Strategic Litigation*, INNOCENCE PROJECT, <http://www.innocenceproject.org/free-innocent/strategic-litigation> (last visited Oct. 9, 2015). The Strategic Litigation Unit employs multiple strategies to achieve this goal, including: the filing of amicus briefs in appropriate cases; consulting and supporting trial attorneys across the country; direct litigation on behalf of individuals at all stages of litigation; training attorneys and judges; and, effectuating change through legislation and policy. See *id.*

41 Mississippi has had an unusually high incidence of bite mark convictions--and post-conviction litigation concerning those convictions-- because state prosecutors used one of its most aggressive practitioners, Dr. Michael West, for two decades, beginning in the mid-1980s. See Radley Balko, *The Bite-Marks Men: Mississippi's Criminal Forensics Disaster*, SLATE (Feb. 20, 2008), http://www.slate.com/articles/news_and_politics/jurisprudence/2008/02/the_bitemarks_men.html.

42 NAS REPORT, *supra* note 4, at 4; see also WILLAM J. STUNTZ, THE COLLAPSE OF AMERICAN CRIMINAL JUSTICE (2011); Craig M. Cooley & Gabrielle S. Oberfield, *Increasing Forensic **Evidence's** Reliability and Minimizing Wrongful Convictions: Applying Daubert Isn't the Only Problem*, 43 TULSA L. REV. 285 (2007); Keith A. Findley, *Judicial Gatekeeping of Suspect **Evidence**: Due Process and Evidentiary Rules in the Age of Innocence*, 47 GA. L. REV. 723 (2013); Sandra G. Thompson, *Judicial Gatekeeping of Police-Generated Witness Testimony*, 102 J. CRIM. L. & CRIMINOLOGY 329 (2012). Scholars have now by and large turned their attention away from critiquing this type of **evidence**-- eyewitness identification, for example--and toward a new type of **evidence**--so-called “second generation” **evidence**, which includes

database-driven techniques such as location tracking, biometrics, and digital forensics--to warn and argue that "the adversarial process, as currently conceived, is ill-suited to ensuring the integrity of high-tech **evidence**." Erin Murphy, *The Mismatch Between Twenty-First Century Forensic Evidence and Our Antiquated Criminal Justice System*, 87 S. CALIF. L. REV. 633, 633 (2015); see also Munia Jabbar, *Overcoming Daubert's Shortcomings in Criminal Trials: Making the Error Rate the Primary Factor in Daubert's Validity Inquiry*, 85 N.Y.U. L. REV. 2034 (2010); Jennifer E. Laurin, *Remapping the Path Forward: Toward a Systemic Review of Forensic Science Reform and Oversight*, 91 TEX. L. REV. 1051 (2013); Erin Murphy, *Inferences, Arguments, and Second Generation Forensic Evidence*, 59 HASTINGS L.J. 1047 (2008); Erin Murphy, *The New Forensics: Criminal Justice, False Certainty, and the Second Generation of Scientific Evidence*, 95 CALIF. L. REV. 721 (2007); Hon. Donald E. Shelton, *Twenty-First Century Challenges for Trial Judges in Criminal Cases: Where the "Polybutadiene" Meets the "Bitumen,"* 18 WIDENER L. REV. 309 (2009).

43 See *supra* note 41 and accompanying text.

44 See, e.g., Erica Beecher-Monas, *Reality Bites: The Illusion of Science in Bite-Mark Evidence*, 30 CARDOZO L. REV. 1369 (2009); D. Michael Risinger, *Navigating Expert Reliability: Are Criminal Standards of Certainty Being Left on the Dock?* 64 ALB. L. REV. 99 (2000) [hereinafter Risinger, *Navigating Expert Reliability*]; Michael J. Saks & Jonathan J. Koehler, *The Individualization Fallacy in Forensic Science Evidence*, 61 VAND. L. REV. 199 (2008). Attention has now also turned to the practicalities of dealing prospectively with certain types of this forensic **evidence**. See e.g., Findley, *supra* note 41.

45 Together, the University of Michigan Law School and the Northwestern University School of Law manage the National Registry of Exonerations, which lists more than 1,000 exonerations since 1989. See *National Registry of Exonerations*, UNIV. OF MICHIGAN LAW SCHOOL, <http://www.law.umich.edu/special/exoneration/Pages/about.aspx> (last visited Oct. 9, 2015).

46 See *infra* Part I.B.

47 See Spencer S. Hsu, *Justice Dept., FBI to Review Use of Forensic Evidence in Thousands of Cases*, WASH. POST (July 10, 2012), https://www.washingtonpost.com/local/crime/justice-dept-fbi-to-review-use-of-forensic-evidence-in-thousands-of-cases/2012/07/10/gJQAT6DlbW_story.html [hereinafter Hsu, *Justice Dept., FBI to Review Use of Forensic Evidence in Thousands of Cases*].

48 See *infra* Part I.C.

49 There have been some isolated and sporadic efforts at the state level to address the problem we identify here. See, e.g., *State v. Henderson*, 27 A.3d 872 (N.J. 2011) (revising the standards for evaluating eyewitness identification testimony so that they more closely track social science findings on reliability); *State v. Lawson*, 291 P.3d 673, 685 (Or. 2012) (finding that "the scientific knowledge and empirical research concerning eyewitness perception and memory has progressed sufficiently to warrant taking judicial notice of ... [them] in determining the effectiveness of our existing test for the admission of eyewitness identification **evidence**").

50 Sarah L. Cooper, *The Collision of Law and Science: American Court Responses to Developments in Forensic Science*, 33 PACE L. REV. 234, 237 (2012).

51 See *id.* at 238.

52 See *id.* (internal citations omitted). In contrast, however, stands some recent Fourth Amendment jurisprudence. Take, for example, *United States v. Jones*, 132 S. Ct. 945 (2011), which considered whether federal law enforcement's attachment of a GPS device to a drug suspect's vehicle constituted a search under the Fourth Amendment. In holding that the aforementioned actions did, in fact, constitute a search, the Court discussed one's right to privacy--previously considered very limited when one was out in public--in an era of secret, electronic monitoring. See *id.*

53 See Cooper, *supra* note 49, at 238 (internal citations omitted).

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE..., 4 Va. J. Crim. L. 1

- 54 See, e.g., Radley Balko, *How the Courts Trap People Who Were Convicted by Bad Forensics*, WASH. POST (Nov. 17, 2014), <http://www.washingtonpost.com/news/thewatch/wp/2014/11/17/how-the-courts-trap-people-who-were-convicted-by-bad-forensics/>.
- 55 See Norah Rudin & Keith Inman, *The Shifty Paradigm, Part I: Who Gets to Define the Practice of Forensic Science?*, 4 CAC NEWS 13 (2005), available at <http://www.forensicdna.com/assets/4thq05.pdf>.
- 56 Koehler & Saks, *The Coming Paradigm Shift in Forensic Identification Science*, *supra* note 3.
- 57 *Id.* at 892. To make traditional forensic individualization sciences fit the new paradigm--and, as a consequence, achieve some level of otherwise lacking scientific rigor--Saks and Koehler argue that DNA typing should be used as a model, noting that: DNA typing technology was an application of knowledge derived from core scientific disciplines ... [and] provided a stable structure for future empirical work Second, the courts and scientists scrutinized applications of the technology in individual cases. As a result, early, unscientific practices were rooted out. Third, DNA typing offered data-based, probabilistic assessments of the meaning of evidentiary “matches.” This practice represented an advance over potentially misleading match/no-match claims associated with other forensic identification sciences.
- Id.* at 893. Saks and Koehler's suggestion is consistent with the NAS Report's findings that “with the exception of nuclear DNA analysis, however, 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. In terms of scientific basis, the analytically based disciplines generally hold a notable edge over disciplines based on expert interpretation.” NAS REPORT, *supra* note 4, at 100.
- 58 The Saks and Koehler article was quite controversial. Norah Rudin and Keith Inman's responsive article, *The Shifty Paradigm, Part I: Who Gets to Define the Practice of Forensic Science?*--while agreeing with Saks and Koehler that DNA science has raised the bar for other forensic disciplines--nonetheless argues that there are core differences between other types of forensic evidence and the access to source populations that make application of DNA-typing models possible. See Rudin & Inman, *supra* note 54; see also Norah Rudin & Keith Inman, *The Shifty Paradigm, Part II: Errors and Lies and Fraud, Oh My!*, 4 CAC NEWS 16 (2006), available at <http://www.forensicdna.com/assets/1stq06.pdf>.
- 59 See, e.g., Samuel R. Gross et al., *Exonerations in the United States 1989 Through 2003*, 95 J. CRIM. L. & CRIMINOLOGY 523 (2005); D. Michael Risinger, *Innocents Convicted: An Empirically Justified Factual Wrongful Conviction Rate*, 97 J. CRIM. L. & CRIMINOLOGY 762 (2007).
- 60 Findley, *supra* note 41, at 727.
- 61 Thompson, *supra* note 41, at 330.
- 62 *Manson v. Brathwaite*, 432 U.S. 98, 110-14 (1977). The Court offered five factors--a list it intended to be non-exclusive--for courts to consider, including: “the opportunity of the witness to view the criminal at the time of the crime, the witness' degree of attention, the accuracy of the witness' prior description of the criminal, the level of certainty demonstrated by the witness at the confrontation, and the length of time between the crime and the confrontation.” *Manson*, 432 U.S. at 114 (citing *Neil v. Biggers*, 409 U.S. 188, 199-200 (1972)). The *Manson* test has been undermined by scientific research that courts have called a “near perfect scientific consensus” that “eyewitness identifications are potentially unreliable in a variety of ways unknown to the average juror” *State v. Guilbert*, 49 A.3d 705, 720-21 (Conn. 2012); see also *State v. Henderson*, 27 A.3d 872, 878 (N.J. 2011) (same); *State v. Lawson*, 291 P.3d 673, 690 n.5 (Or. 2012) (noting frequency of misidentification). See also NAT'L RESEARCH COUNCIL OF THE NAT'L ACADEMIES, IDENTIFYING THE CULPRIT: ASSESSING EYEWITNESS IDENTIFICATION 13 (2014), available at https://public.psych.iastate.edu/glwells/NAS_Eyewitness_ID_Report.pdf (noting that *Manson* “was not based on much of the research conducted by scientists on visual perception, memory, and eyewitness identification, and ... fails to include important advances”).
- 63 *Henderson*, 27 A.3d at 872.
- 64 *Id.* at 884-85.

65 *Id.* at 919-24.

66 *Id.* at 919.

67 *Lawson*, 291 P.3d at 673.

68 *Id.* at 690-97.

69 *Id.* at 697.

70 Since 1989, there have been more than 300 documented exonerations based on post-conviction DNA testing. *See DNA Exonerations Nationwide*, INNOCENCE PROJECT, (Sep. 3, 2015, 12:30 PM) <http://www.innocenceproject.org/free-innocent/improve-the-law/fact-sheets/dna-exonerations-nationwide>.

71 *See* Brandon L. Garrett, *Judging Innocence*, 108 COLUM. L. REV. 55, 59-60, 76 (2007).

72 *See* Garrett & Neufeld, *supra* note 10, at 2.

73 Garrett, *supra* note 70, at 81.

74 *Id.*

75 *Id.* at 82. For example, Garrett reports “[a] preliminary review of serological testimony during these exonerees' trials disclosed that more than half involved improper testimony by forensic examiners.” *Id.*

76 *Id.* at 83.

77 An appointed task force created during an inspector general's investigation of misconduct at the FBI crime lab in the 1990s undertook the investigation. Spencer S. Hsu, *Convicted Defendants Left Uninformed of Forensic Flaws Found by Justice Department*, WASH. POST (April 16, 2012), https://www.washingtonpost.com/local/crime/convicted-defendants-left-uninformed-of-forensic-flaws-found-by-justicedept/2012/04/16/gIQAWTcgMT_story.html. The inquiry lasted nine years and ended in 2004. *Id.*

78 *Id.*

79 *Id.*

80 Spencer S. Hsu, *Review of FBI Forensics Does Not Extend to Federally Trained State, Local Examiners*, WASH. POST (Dec. 22, 2012), https://www.washingtonpost.com/local/crime/review-of-fbi-forensics-does-not-extend-to-federally-trained-state-local-examiners/2012/12/22/b7ef9c2e-4965-11e2-ad54-580638ede391_story.html. *See also* Hsu, *Convicted Defendants Left Uninformed of Forensic Flaws Found by Justice Department*, *supra* note 76, which states that:

The Washington Post worked with the nonprofit National Whistleblowers Center, which had obtained dozens of boxes of task force documents through a years-long Freedom of Information Act fight. Task force documents identifying the scientific reviews of problem cases generally did not contain the names of the defendants. Piecing together case numbers and other bits of information from more than 10,000 pages of documents, The Post found more than 250 cases in which a scientific review was completed. Available records did not allow the identification of defendants in roughly 100 of those cases. Records of an unknown number of other questioned cases handled by federal prosecutors have yet to be released by the government.
Id.

81 There is considerable **evidence** that the FBI trained all examiners how to testify and exaggerate their findings beyond the limits of science. *See* Clyde Haberman, *DNA Analysis Exposes Flaws in an Inexact Forensic Science*, N.Y. TIMES (May 18, 2014), http://www.nytimes.com/2014/05/19/us/dna-analysis-exposes-an-inexact-forensic-science.html?_r=0 (“A forensics expert who used to work in the federal lab, Max M. Houck, told [The New York Times] Retro Report that there was ‘absolutely a disconnect between what I could say as a scientist and what the prosecutors, or the defense attorneys, wanted me to say.’”); *see also* Spencer S. Hsu, *FBI Lab's Woes Cast Growing Shadow*, INDEPENDENT (Dec. 23, 2012), <http://www.independent.co.uk/>

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE..., 4 Va. J. Crim. L. 1

news/world/americas/fbi-labs-woes-cast-a-growing-shadow-8430348.html (“[A]bout three dozen FBI agents trained 600 to 1,000 state and local examiners to apply the same standards that have proved problematic.”).

82 Amanda L. Myers, *Once Key in Some Cases, Bite Mark Evidence Now Derided as Unreliable*, DENVER POST (June 17, 2013), http://www.denverpost.com/ci_23474835/once-key-some-cases-bite-mark-evidence-now [hereinafter Myers, *Once Key in Some Cases, Bite Mark Evidence Now Derided as Unreliable*].

83 Amanda L. Myers, *Men Wrongly Convicted or Arrested on Bite Evidence*, YAHOO! NEWS (June 16, 2013), <http://news.yahoo.com/men-wrongly-convicted-arrested-bite-evidence-150610286.html> [hereinafter Myers, *Men Wrongly Convicted or Arrested on Bite Evidence*].

84 This statement is based primarily on the authors' experience, both of whom have represented thousands of criminal defendants in previous careers as public defenders with the Public Defender Service of Washington, D.C and The Bronx Defenders.

85 See, e.g., Paul C. Giannelli, *Daubert and Forensic Science: The Pitfalls of Law Enforcement Control of Scientific Research*, 2011 U. ILL. L. REV. 53 (2011) [hereinafter Giannelli, *Daubert and Forensic Science*]; David H. Kaye, *The Good, the Bad, and the Ugly: The NAS Report on Strengthening Forensic Science in America*, 50 SCI. & JUST. 8 (2010); Kenneth E. Melson, *Embracing the Path Forward: The Journey to Justice Continues*, 36 NEW ENG. J. ON CRIM. & CIV. CONFINEMENT 197 (2010).

86 NAS REPORT, *supra* note 4, at 4.

87 See *id.* at 14-15.

88 *Id.* at 15.

89 *Id.* at 14-15.

90 See *supra* note 57 and accompanying text.

91 NAS REPORT, *supra* note 4, at 42-43.

92 See e.g., Giannelli, *Daubert and Forensic Science*, *supra* note 83, at 57; Paul C. Giannelli, “*Ballistics*” *Redux*, 28 CRIM. JUST. 47, 48 (2013). Jennifer L. Mnookin et al., *The Need for a Research Culture in the Forensic Sciences*, 58 UCLA L. REV. 725, 729-32 (2011); William A. Tobin & Peter J. Blau, *Hypothesis Testing of the Critical Underlying Premise of Discernible Uniqueness in Firearms-Toolmarks Forensic Practice*, 53 JURIMETERICS J. 121, 141-42 (2013).

93 See NAT'L SCI. & TECH. COUNCIL, COMM. ON SCI., SUBCOMM. ON FORENSIC SCI., STRENGTHENING THE FORENSIC SCIENCES (2014), available at https://www.whitehouse.gov/sites/default/files/microsites/ostp/NSTC/strengthening_the_forensic_sciences_may_-_2014.pdf.

94 This portion of the article would not have been possible without the expertise of Sarah Chu, Senior Forensic Policy Advocacy at the Innocence Project, and her encyclopedic knowledge of the federal forensic science landscape.

95 See Press Release, Dep't of Justice, Department of Justice and National Institute of Standards and Technology Announce Launch of National Commission on Forensic Science (Feb. 15, 2013), available at <http://nist.gov/oles/doj-nist-forensic-science021513.cfm>.

96 DEPT OF JUST. & NAT'L INST. OF STANDARDS & TECH., MEMORANDUM OF UNDERSTANDING BETWEEN THE DEPARTMENT OF JUSTICE AND THE NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY IN SUPPORT OF THE NATIONAL COMMISSION ON FORENSIC SCIENCE AND THE ORGANIZATION OF SCIENTIFIC AREA COMMITTEES 1 (2015).

97 The NCFs Charter provides that:

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE..., 4 Va. J. Crim. L. 1

The objectives and scope of activities of the Commission are to provide recommendations and advice to the Department of Justice (DOJ) concerning national methods and strategies for: strengthening the validity and reliability of the forensic sciences (including medico-legal death investigation); enhancing quality assurance and quality control in forensic science laboratories and units; identifying and recommending scientific guidance and protocols for **evidence** seizure, testing, analysis, and reporting by forensic science laboratories and units; and identifying and assessing other needs of the forensic science communities to strengthen their disciplines and meet the increasing demands generated by the criminal and civil justice systems at all levels of government. In accomplishing these objectives, the Commission may not develop or recommend guidance regarding digital **evidence**.

DEPT OF JUSTICE, NAT'L COMM'N ON FORENSIC SCI. CHARTER 1 (Mar. 18, 2013), *available at* <http://www.justice.gov/ncfs/docs/ncfs-charter.pdf>.

98 *See* DEPT OF JUST. & NAT'L INST. OF STANDARDS AND TECH., *supra* note 94, at 2.

99 Mark D. Storolow, *MOU on Forensic Science Between NIST and US DOJ*, NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY (Jun. 11, 2013), <http://www.nist.gov/director/vcat/upload/Storolow-Forensic-Science-MOU.pdf>; *see* DEPT OF JUST. & NAT'L INST. OF STANDARDS & TECH., *supra* note 94, at 3.

100 DEPT OF JUST. & NAT'L INST. OF STANDARDS & TECH., *supra* note 94, at 3; *see also* Storolow, *supra* note 97.

101 *See* Press Release, Dep't of Justice, U.S. Departments of Justice and Commerce Name Experts to First-ever National Commission on Forensic Science (Jan. 10, 2014), *available at* <http://www.justice.gov/opa/pr/2014/January/14-at-029.html>.

102 *See* NAT'L ASSOC. OF FORENSIC SCI., MEETING SUMMARY 11-16 (Feb. 3-4, 2014), *available at* <http://www.justice.gov/ncfs/docs/meeting-sum.pdf> [hereinafter MEETING SUMMARY].

103 Michael E. Newman, *New NIST Center of Excellence to Improve Statistical Analysis of Forensic Evidence*, NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY (May 26, 2015), <http://nist.gov/forensics/center-excellence-forensic052615.cfm>. The Forensic Science Center of Excellence is led by Iowa State in partnership with Carnegie Mellon University, the University of Virginia, and the University of California, Irvine. *Id.*

104 *See id.*

105 *Id.* The National Institute of Standards and Technology goes on to describe this research agenda as follows: NIST and university scientists working within the new center will develop tools to evaluate pattern and digital **evidence** analysis methods for how well they consider statistical modeling errors and uncertainties in measurement. This will allow forensic scientists to quantify the level of confidence they have in statistical computations made with these methods and the conclusions reached from those analyses.

Id. As discussed in Parts II.A and III.A, courts have allowed experts to express probabilist conclusions in pattern **evidence** techniques, even in the absence of this foundational research, for decades. *See infra* Part II.A, Part III.A.

106 The OSAC was originally titled “guidance groups” in the MOU. *See* NAT'L INST. OF STANDARDS AND TECH., SUMMARY OF THE NIST PROPOSED PLAN FOR THE ORGANIZATION OF SCIENTIFIC AREA COMMITTEES (OSAC) 1 (2014), *available at* <http://nist.gov/forensics/upload/NIST-OSAC-Summary-1-31-14.pdf> [hereinafter SUMMARY OF THE NIST PROPOSED PLAN].

107 SUMMARY OF THE NIST PROPOSED PLAN, *supra* note 104, at 1.

108 Scientific Working Groups (SWG) are groups of “scientific subject-matter experts who collaborate to determine best practices and to develop consensus standards.” *Legacy Scientific Working Groups*, NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY, <http://www.nist.gov/forensics/workgroups.cfm> (last visited Feb. 13, 2016). Previously, the FBI nominally supported SWGs. *See* Barbara Guttman & John M. Butler, *Plan for the Organization of Scientific Area Committees (OSAC)*, NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY, <http://www.nist.gov/forensics/upload/AAFS-NIST-OSAC-presentation-Feb-24-2014.pdf> (last visted Feb. 13, 2016).

- 109 See SUMMARY OF THE NIST PROPOSED PLAN, *supra* note 104, at 1-2. The question surrounding the composition of subcommittees and the level of incorporation of researchers and forensic scientists has elicited a range of responses in the Public Comments on the NIST Notice of Inquiry. See generally NAT'L INST. OF STANDARDS AND TECH., PUBLIC COMMENTS ON NIST NOTICE OF INQUIRY: POSSIBLE MODELS FOR THE ADMINISTRATION AND SUPPORT OF DISCIPLINE SPECIFIC GUIDANCE GROUPS FOR FORENSIC SCIENCE (2013), available at <http://nist.gov/forensics/upload/commentspdf-020714-small-3-2.pdf>.
- 110 SUMMARY OF THE NIST PROPOSED PLAN, *supra* note 104, at 1.
- 111 See *id.* at 1-2.
- 112 See MEETING SUMMARY, *supra* note 99; see also Press Release, Nat'l Inst. of Standards and Tech., 402 Members Named to Forensic Science Standards Organization (Oct. 29, 2014), available at http://www.nist.gov/forensics/osac_102914.cfm.
- 113 For example, the Department of Justice recently announced that it will “conduct a quality assurance review of other forensic science disciplines practiced at the FBI—to determine whether the same kind of ‘testimonial overstatement’ we found during our review of microscopic hair evidence could have crept into other disciplines that rely heavily on human interpretation and where the degree of certainty can be difficult to quantify.” Sally Q. Yates, Deputy Attorney General, Remarks During the 68th Annual Scientific Meeting Hosted by the American Academy of Forensic Science (February 24, 2016), available at <https://www.justice.gov/opa/speech/deputy-attorney-general-sally-q-yates-delivers-remarks-during-68th-annual-scientific>. See also Spencer S. Hsu, *Justice Dept. to Expand Review of FBI Forensic Techniques Beyond Hair Unit*, WASH. POST (Feb. 25, 2016), https://www.washingtonpost.com/local/public-safety/justice-dept-to-expand-review-of-fbi-forensic-techniques-beyond-hair-unit/2016/02/25/5adf0b8c-dbd4-11e5-81ae-7491b9b9e7df_story.html.
- 114 See Spencer S. Hsu, *Justice Dept., FBI to Review Use of Forensic Evidence in Thousands of Cases*, *supra* note 46.
- 115 See Hsu, *Convicted Defendants Left Uninformed of Forensic Flaws Found by Justice Department*, *supra* note 76.
- 116 Certificate of Innocence at 2, *United States v. Tribble*, No. 78 FEL 4160 (D.D.C. Dec. 14, 2012), available at <https://archive.org/stream/539840-tribble#page/n0/mode/2up>.
- 117 *Id.* at 1.
- 118 See Spencer S. Hsu, *D.C. Judge Exonerates Santae Tribble in 1978 Murder, Cites Hair Evidence DNA Test Rejected*, WASH. POST (Dec. 14, 2012), https://www.washingtonpost.com/local/crime/dc-judge-exonerates-santae-tribble-of-1978-murder-based-on-dna-hair-test/2012/12/14/da71ce00-d02c-11e1-b630-190a983a2e0d_story.html.
- 119 See Spencer S. Hsu, *Kirk Odom, Who Served 20 Years for 1981 D.C. Rape, Is Innocent, Prosecutors Say*, WASH. POST (July 10, 2012), https://www.washingtonpost.com/local/crime/2012/07/10/gJQAUjZNbW_story.html.
- 120 See *id.*
- 121 Hsu, *Convicted Defendants Left Uninformed of Forensic Flaws Found by Justice Department*, *supra* note 76.
- 122 See *supra* Part III.B. More specifically, the errors have been identified as follows:
Error Type 1: The examiner stated or implied that the evidentiary hair could be associated with a specific individual to the exclusion of all others.
Error Type 2: The examiner assigned to the positive association a statistical weight or probability or provided a likelihood that the questioned hair originated from a particular source, or an opinion as to the likelihood or rareness of the positive association that could lead the jury to believe that valid statistical weight can be assigned to a microscopic hair association.
Error Type 3: The examiner cites the number of cases or hair analyses worked in the lab and the number of samples from different individuals that could not be distinguished from one another as a predictive value to bolster the conclusion that a hair belongs to a specific individual.

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE..., 4 Va. J. Crim. L. 1

FED. BUREAU OF INVESTIGATION, MICROSCOPIC HAIR COMPARISON ANALYSIS AGREEMENT 1 (Nov. 9, 2012), available at <http://www.mtacdl.org/attachments/CPE/Nelson/FBIlimitsofScience%20MicroscopicHairComparison.pdf>.

- 123 See Hsu, *Convicted Defendants Left Uninformed of Forensic Flaws Found by Justice Department*, supra note 76. In his article, Hsu noted that:
The [Washington] Post worked with the nonprofit National Whistleblowers Center, which had obtained dozens of boxes of task force documents through a years-long Freedom of Information Act fight. Task force documents identifying the scientific reviews of problem cases generally did not contain the names of the defendants. Piecing together case numbers and other bits of information from more than 10,000 pages of documents, The Post found more than 250 cases in which a scientific review was completed. Available records did not allow the identification of defendants in roughly 100 of those cases. Records of an unknown number of other questioned cases handled by federal prosecutors have yet to be released by the government.
Id.
- 124 See *id.*
- 125 See Spencer S. Hsu, *U.S. Reviewing 27 Death Penalty Convictions for FBI Forensic Testimony Errors*, WASH. POST (July 17, 2013), https://www.washingtonpost.com/local/crime/us-reviewing-27-death-penalty-convictions-for-fbi-forensic-testimony-errors/2013/07/17/6c75a0a4-bd9b-11e2-89c9-3be8095fe767_story.html. According to the article, “[t]he death row cases are among the first 120 convictions identified as potentially problematic among more than 21,700 FBI Laboratory files being examined.” *Id.*
- 126 For an excellent discussion of this case and the appellate treatment of it, see Andrew Cohen, *A Ghost of Mississippi: The Willie Manning Capital Case*, ATLANTIC (May 2, 2013), <http://www.theatlantic.com/national/archive/2013/05/a-ghost-ofmississippi-the-willie-manning-capital-case/275442/>.
- 127 See Order, *Manning v. State*, No. 2013-DR-00491-SCT (Miss. Apr. 25, 2013), available at <http://courts.ms.gov/Images/Opinions/183549.pdf>.
- 128 Letter from John Crabb, Jr., Special Counsel, U.S. Dep't of Justice, to Deforest R. Allgood, Oktibbeha Cnty. Mississippi District Attorney's Office (May 2, 2013), available at <http://www.scribd.com/doc/139768317/DOJ-Letter-in-Manning-Case-May-2nd>. More specifically, the letter stated that:
We have determined that the microscopic hair comparison analysis testimony or lab analysis report presented in this case included statements that exceeded the limits of science, and was, therefore, invalid. While this case did not involve a positive association of an evidentiary hair to an individual, the examiner stated or implied in a general explanation of microscopic hair comparison analysis that a questioned hair could be associated with a specific individual to the exclusion of all others - this type of testimony exceeded the limits of the science. The examiner also assigned a statistical weight or probability or provided a likelihood that, through microscopic hair comparison analysis, the examiner could determine that a questioned hair originated from a particular source, or an opinion as to the likelihood or rareness of a positive association that could lead the jury to believe that valid statistical weight can be assigned to a microscopic hair association - this type of testimony exceeded the limits of the science.
Id.
- 129 See Order, *Manning v. State*, 112 So. 3d 1082 (Miss. 2013) (No. 95-DP-00066-SCT).
- 130 See Order, *Manning v. State*, 119 So. 3d 293 (Miss. 2013) (No. 2013-DR-00491-SCT).
- 131 See, e.g., *In re Richards*, 289 P.3d 860 (2012). *California Lawyer* determined that *In re Richards* was the worst state court decision of 2012 because it created a “distinction between the testimony of experts and the testimony of laypersons in applying the protections against false evidence in Penal Code section 1473(b)” and thus “create[d] a substantial obstacle to correcting what the California Commission on the Fair Administration of Justice identified as the second-most-common factor contributing to wrongful convictions: erroneous scientific evidence.” Gerald F. Uelmen,

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE..., 4 Va. J. Crim. L. 1

New Balance at the California Supreme Court, CAL. LAWYER (Aug. 2013), http://www.callawyer.com/Clstory.cfm?eid=930177&wteid=930177_New_Balance_at_the_California_Supreme_Court.

- 132 Compare *Ex parte Henderson*, 246 S.W.3d 690 (Tex. Crim. App. 2007) (granting state habeas relief to a woman who was previously convicted of killing a baby in her care because biomechanical evidence showed that the death could have been the result of an accident rather than an intentional act), with *Ex parte Robbins*, 360 S.W.3d 446 (Tex. Crim. App. 2011) (denying state habeas relief despite the testimony of several medical examiners, including the one who performed the original autopsy, which concluded that the cause of death was “undetermined” rather than “homicide”).
- 133 See Tex. H.B. 1068, 79th Leg., R.S. (Tex. 2005).
- 134 TEX. CODE CRIM. PROC. ANN. art. 38.01 § 4(a)(3) (West 2015). The Commission has nine members who are appointed by the Governor. See *id.* at § 3(a).
- 135 See David Grann, *Trial by Fire: Did Texas Execute an Innocent Man?*, NEW YORKER (Sept. 7, 2009), <http://www.newyorker.com/magazine/2009/09/07/trial-by-fire>.
- 136 See Possley, *supra* note 30.
- 137 See Brandi Grissom, *New Head of Forensic Science Panel Takes on Arson Case*, TEXAS TRIB. (July 22, 2011), <http://www.texastribune.org/2011/07/22/new-head-of-forensic-science-panel-takes-on-arson/>.
- 138 See Possley, *supra* note 30. Thereafter, the Texas Attorney General issued an opinion letter that limited the Commission's jurisdiction, including its investigation into the *Willingham* case, by prohibiting consideration of any cases where evidence was submitted before September 1, 2005. *Willingham* was executed in 2004. See *Whether the Science Advisory Workgroup of the State Fire Marshal's Office Has Authority to Review Prior Arson Investigations*, Op. Tex. Att'y Gen. GA-1048 (Apr. 4, 2014), available at <https://www.texasattorneygeneral.gov/opinions/opinions/50abbott/op/2014/pdf/ga1048.pdf>.
- 139 Texas Senate Bill 1611, known as the “Michael Morton Act,” broadens defendants' access to evidence that could prove innocence. Tex. S.B. 1611, 83rd Leg., R.S. (Tex. 2013) (enacted). Senate Bill 344 allows a defendant to challenge a conviction that was gained through forensic techniques that have since been discredited by modern science. Tex. S.B. 344, 83rd Leg., R.S. (Tex. 2013) (enacted). This bill responds to documented problems in old arson cases where junk science was used as evidence. *Id.*
- 140 See *id.* The statute is, in part, an answer to the *Willingham* case, but it is also the result of two inconsistent rulings. In *Ex parte Henderson*, 246 S.W.3d 690 (Tex. Crim. App. 2007), the Texas Court of Criminal Appeals granted habeas relief to a woman who was previously convicted of killing a baby in her care because biomechanical evidence showed that the death could have been the result of an accident rather than an intentional act. In *Ex Parte Robbins*, 360 S.W.3d 446, 453 (Tex. Crim. App. 2011), the Texas Court of Criminal Appeals denied habeas relief to a man who was convicted of killing a toddler in his care despite the testimony of several medical examiners, including the medical examiner who performed the original autopsy, which concluded that the cause of death was “undetermined” rather than “homicide.” The court concluded that Robbins failed to show that the testimony given by the medical examiner during the trial was false. *Id.* at 460. Moreover, an actual innocence claim required Robbins to show “by clear and convincing evidence that no reasonable juror would have convicted him in light of” the medical examiner's recantation, and Robbins failed to do so. *Id.* at 459. The new statute lowers the burden from a “clear and convincing” standard to a “preponderance of the evidence.” *Id.* at 459; see also Tex. S.B. 344, 83rd Leg., R.S. (Tex. 2013) (enacted). On November 26, 2014, Robbins was the first petitioner to be granted relief under the statute's new terms. See *Ex parte Robbins*, No. WR-73, 484-02 (Tex. Crim App. Nov. 25, 2014).
- 141 See S.B. 1058, 113th Cong. (Cal. 2014) (enacted).
- 142 See *id.*
- 143 Garrett & Neufeld, *supra* note 10, at 90.

- 144 One of the more egregious, but emblematic, examples includes the following case. In Eddie Lee Howard's death penalty conviction, the Mississippi Supreme Court was confronted with a host of valid claims about Dr. Michael West--the forensic odontologist's--malfeasance, including instances where Dr. West had misidentified bite marks in other cases. *See Howard v. State*, 945 So. 2d 326, 352 (Miss. 2006). With regard to the admission of his testimony in Howard's case, though, the court wrote:
In support of his post-conviction claim, Howard has offered numerous expert affidavits and other documents which attack Dr. West, his testimony, and bite mark **evidence** in general. These affidavits and other documents point out how many times Dr. West has been proven wrong and they discuss how unscientific his methods are. One affidavit even states that Dr. West made a misdiagnosis in Howard's case, but, it does not go on and opine that Howard did not bite Kemp. Just because Dr. West has been wrong a lot, does not mean, without something more, that he was wrong here.
Id.
- 145 Garrett & Neufeld, *supra* note 10, at 90-92.
- 146 *See infra* notes 227, 311 and accompanying text.
- 147 *Frye v. United States*, 293 F. 1013, 1014 (D.C. Cir. 1923) (requiring the proponent of novel scientific **evidence** to bear the burden of demonstrating that the **evidence** has “gained general acceptance in the particular field in which it belongs”).
- 148 *Daubert v. Merrill Dow Pharm.*, 509 U.S. 579 (1993).
- 149 *See infra* pp. 41-45 (deconstructing precedent-establishing cases in the field of forensic odontology).
- 150 *See infra* pp. 58-60 (describing the narrow definitions of the relevant scientific community utilized in bite mark jurisprudence).
- 151 Samuel R. Gross & Jennifer L. Mnookin, *Expert Information and Expert Evidence: A Preliminary Taxonomy*, 34 SETON HALL L. REV. 141, 169 (2003) (“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.”).
- 152 126 Cal. Rptr. 350 (Cal. Ct. App. 1975).
- 153 *Id.*
- 154 *Frye*, 293 F. at 1014.
- 155 *Marx*, 126 Cal. Rptr. at 352-53, 355-57.
- 156 Courts, including the *Marx* court, have conflated the identification of human remains through dental records with bite mark identification. The former is a well-established, relatively non-controversial technique; bite mark analysis is an entirely different discipline, relying on untested assumptions and the interpretation of injuries in human flesh. Nonetheless, courts have often treated the disciplines as essentially interchangeable, further insulating bite mark **evidence** from judicial scrutiny. *See, e.g., Handley v. State*, 515 So. 2d 121, 129 (Ala. Crim. App. 1987); *Marx*, 126 Cal. Rptr. 350; *People v. Milone*, 396 N.E.2d 1350 (Ill. 1976); *People v. Middleton*, 429 N.E.2d 100, 103 (N.Y. 1981); *State v. Jones*, 259 S.E.2d 120, 124 (S.C. 1979).
- 157 Risinger, *Navigating Expert Reliability*, *supra* note 43, at 138 (“*Marx* came to be read as a global warrant to admit bite mark identification **evidence** whenever a person displaying apparent credentials chose to testify to an identification.”).
- 158 For the purposes of this article, we are using the terms “science” and “hypothesis” as the Supreme Court defined them in *Daubert v. Merrell Dow Pharm.*, 509 U.S. 579 (1993). As Professor Edward Imwinkelried has explained, “the Court defined science as the process of formulating hypotheses about phenomena in the corporeal world and then engaging in experimentation or observation to falsify or validate the hypothesis. The Court decreed that to qualify his or her proffered testimony as ‘scientific ... knowledge,’ the proponent must lay a foundation establishing the empirical validation of the expert's underlying hypothesis.” Edward J. Imwinkelried, *Should the Courts Incorporate a Best Evidence Rule into the Standard Determining the Admissibility of Scientific Testimony?: Enough Is Enough Even When It Is Not the Best*, 50 CASE W. RES. L.

REV. 19, 22 (1999) (internal citations omitted). What we have defined as the two underlying “hypotheses” of bite mark **evidence** and hair comparison **evidence** (*see infra* notes 201, 233 and accompanying text) are the authors' construct, which we have advanced here to explore whether the interpretations of the data collected and proffered as **evidence** of the identity of a suspect have been scientifically validated. Scientific “validation” answers the question of “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: CASES AND MATERIALS 60 (6th ed. 2006). Although we are using the term “hypotheses,” this might also be conceived simply as two steps in the forensic identification process. The first step of the process establishes the “reliability” of the technique by determining whether the questioned item of **evidence** and the exemplar from a known source do, in fact, share sufficiently similar characteristics to make an association. *See Saks & Koehler, The Individualization Fallacy in Forensic Science Evidence*, *supra* note 43, at 199 (describing the two fundamental steps of forensic identification science); *see also* DAVID A. SCHUM, **EVIDENTIAL FOUNDATIONS OF PROBABILISTIC REASONING** (1994) (offering a general theory of **evidence** as it is understood and applied across disciplines, including law). The second step looks at the “diagnosticity” of the **evidence** by assessing the meaning of the match: “What is the probability that the questioned and the known originated from the same source?” Saks & Koehler, *The Individualization Fallacy in Forensic Science Evidence*, *supra* note 43, at 199. Whether referred to as hypotheses for simplicity or steps, both parts of the bite mark analysis process need to be validated scientifically. *See* Peter J. Neufeld & Neville Colman, *When Science Takes the Witness Stand*, 262 SCI. AM. 46, 49-52 (1990) (describing the importance of validity in the first step of forensic science--reliability of the pattern match--and stressing the importance of validity in the second step--diagnosticity--by examining the “data and assumptions on which forensic laboratories have been relying to estimate the rarity”).

159 Robust reporting of error rates in the field does not exist, and at least one commentator has suggested an affirmative reason for that. *See* Risinger, *Navigating Expert Reliability*, *supra* note 43, at 142 (“[B]ite mark experts have benefited from their ability ... to do few proficiency studies and to keep secret the results of such proficiency studies.”); *see also* C. Michael Bowers, *Problem-Based Analysis of Bitemark Misidentifications: The Role of DNA*, 159 FORENSIC SCI. INT'L S104, S106-S107 (2006). Even the results of controlled studies have been disturbing. At a 1999 American Board of Forensic Odontology bite mark workshop, “ABFO diplomats attempted to match four bitemarks to seven dental models [and] found 63.5% false positives.” Bowers, *supra*, at S106. A 2001 study of “bites made in pig skin, ‘widely accepted as an accurate analogue of human skin,’” resulted in 11.9 to 22.0 percent “false positive identifications ... for various groups of forensic odontologists.” *Id.*

160 *See* NAS REPORT, *supra* note 4, at 173-76 (assessing the current state of bite mark analysis).

161 Take, for example, Mississippi appellate courts--the same courts that affirmed, among others, Levon Brooks' and Eddie Lee Howard's convictions, both of which were based on bite mark **evidence**. *See Howard v. State*, 853 So. 2d 781, 784-85 (Miss. 2003); *Brooks v. State*, 748 So. 2d 736, 739 (Miss. 1999); *Levon Brooks*, *supra* note 20. The Mississippi Supreme Court has spent considerable time discussing the merits--or lack thereof--of expert testimony in a number of different contexts, including the cause of plaintiff's need for hip-replacement surgery where the expert lacked experience and training in orthopedics. *See, e.g., Bailey Lumber & Supply Co. v. Robinson*, 98 So. 3d 986 (Miss. 2012) (“[T]he expert opinion of a doctor as to causation must be expressed in terms of medical probabilities as opposed to possibilities.”); *Univ. of Miss. Med. Center v. Lanier*, 97 So. 3d 1197 (Miss. 2012) (noting that (1) “when the reliability of an expert's opinion is attacked with credible **evidence** that the opinion is not accepted within the scientific community, the proponent of the opinion under the attack should provide at least a minimal defense supporting the reliability of the opinion,” and (2) that an “offered opinion that has been contradicted by published and peer-reviewed data, however, must be supported by some **evidence** of support and acceptance in the scientific community”); *Sherwin Williams Co. v. Gaines*, 75 So. 3d 41 (Miss. 2011) (holding as error the admission of expert testimony on the present value of the utility's future cash flow where the expert acknowledged that his valuation was merely his opinion with no supporting methodology); *McKee v. Bowers Window & Door Co.*, 64 So. 3d 926 (Miss. 2011) (denying expert qualification by distinguishing potential expertise as a general contractor from that of a storm window installation specialist); *Dedeaux Util. Co. v. Gulfport*, 63 So. 3d 514 (Miss. 2011) (engaging in close scrutiny of, among other things, utility cashflow and the relative valuation of storm windows); *Patterson v. Tibbs*, 60 So. 3d 742 (Miss. 2011); (holding that the testimony of the plaintiff's three experts should have been excluded because the experts' opinion that exposure to lead paint caused plaintiff's brain injury was not based on “any scientific authority that acute, asymptomatic ingestion of lead could lead to the alleged injuries,” and

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE..., 4 Va. J. Crim. L. 1

likewise reminding trial judges that their gatekeeping duty under *Daubert* “includes making sure that the opinions themselves are based on sufficient facts or data and are the product or reliable principles and methods”).

162 See *People v. Marx*, 126 Cal. Rptr. 350, 353-55 (Cal. Ct. App. 1975).

163 *Id.* at 353.

164 *Id.*

165 *Id.* at 354.

166 *Id.* at 355.

167 *Id.*

168 *Id.* at 356.

169 *Id.*

170 *Id.* at 356 n.14.

171 See Myers, *Men Wrongly Convicted or Arrested on Bite Evidence*, *supra* note 82 (detailing twenty-four cases).

172 See, e.g., *Davasher v. State*, 823 S.W.2d 863 (Ark. 1992) (“The State requested this evidence to prove by scientific testimony that a wound located on Davasher’s leg was a bite inflicted by the [victim’s] dog, Scooter. Dr. Richard Glass, a forensic odontologist, was allowed to testify that he could not rule Scooter out as the dog that bit Davasher.”).

173 See, e.g., *State v. Armstrong*, 179 S.E.2d 870 (W. Va. 1988) (“[A]n examination of each tooth indicates an exact, perfect match between the appellant’s teeth and the bitemark pattern on the paper towel, with no incompatibility. Dr. Sopher therefore concluded with a reasonable degree of dental certainty that ‘the bite-mark pattern in the towel is that of the teeth of Keith Armstrong, to the exclusion of all other individuals.’”).

174 *Handley v. State*, 515 So. 2d 121, 129 (Ala. Crim. App. 1987); *State v. Kleypas*, 602 S.W.2d 863, 869 (Mo. Ct. App. 1980) (“[A] scientific principle ... may or may not be involved in such a [bite mark] comparison.”); *Kennedy v. State*, 640 P.2d 971, 978 (Okla. Crim. App. 1982).

175 *People v. Slone*, 76 Cal. App. 3d 611 (Cal. Ct. App. 1978).

176 *Id.* at 624.

177 *Id.*

178 *Id.*

179 The authors are indebted to University of Baltimore Law Professor Colin Starger, who created this graphic using software that he has developed for mapping Supreme Court precedent. See *The Supreme Court Mapping Project*, UNIV. OF BALTIMORE, <http://law.ubalt.edu/faculty/scotus-mapping/> (last visited Oct. 10, 2015).

180 This phenomenon is true not only of bite mark evidence, but also of many other forensic techniques. See, e.g., Jules Epstein, *Preferring the “Wise Man” to Science: The Failure of Courts and Non-Litigation Mechanisms to Demand Validity in Forensic Matching Testimony*, 20 WIDENER L. REV 81, 114 (2014) (noting the National Research Council’s conclusions that available data does not support matching a bullet to a particular “box” of ammunition; that compositional analysis does not support definitive statements about the date of bullet manufacture; and, that detailed distribution of ammunition is such that probabilistic claims that a specific bullet came from a defendant should be avoided); NAT’L RESEARCH COUNCIL OF THE NAT’L ACADEMIES, *FORENSIC ANALYSIS: WEIGHING BULLET LEAD EVIDENCE* 1 (2004), available at http://www.nap.edu/catalog.php?record_id=10924.

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE..., 4 Va. J. Crim. L. 1

- 181 See, e.g., AM. BOARD OF FORENSIC ODONTOLOGY, DIPLOMATES REFERENCE MANUAL (2013) [hereinafter DIPLOMATES REFERENCE MANUAL], available at <http://www.abfo.org/wp-content/uploads/2012/08/ABFO-Reference-Manual-March-2015.pdf>.
- 182 See G.S. Golden, *Standards and Practices for Bite Mark Photography*, 29 J. FORENSIC ODOTOSTOMATOLOGY 29, 29, 32 (2011).
- 183 *Id.* at 29, 32.
- 184 See generally Robert B.J. Dorion, *Transillumination in Bite Mark Evidence*, 32 J. FORENSIC SCI. 690 (1987).
- 185 See ANGI M. CHRISTENSEN ET AL., FORENSIC ANTHROPOLOGY: CURRENT METHODS AND PRACTICE 81 (2014).
- 186 BITEMARK EVIDENCE: A COLOR ATLAS AND TEXT 153 (Robert B.J. Dorion ed., 2d ed. 2011).
- 187 See *id.* at 112.
- 188 See *id.*
- 189 DIPLOMATES MANUAL, *supra* note 181, at 96.
- 190 *Id.* at 101.
- 191 *Id.*
- 192 See Dorion, *supra* note 182, at 690.
- 193 DIPLOMATES MANUAL, *supra* note 181, at 97.
- 194 See *id.* at 96, 98.
- 195 See, e.g., *Daubert v. Merrell Dow Pharm.*, 509 U.S. 579 (1993) (“Expert evidence can be both powerful and quite misleading because of the difficulty in evaluating it.”); *United States v. Frazier*, 387 F.3d 1244, 1263 (11th Cir. 2004) (“Simply put, expert 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.”); NEIL VIDMAR & VALERIE P. HANS, *AMERICAN JURIES: THE VERDICT* 181-82 (2007); B. Michael Dann et. al., *Can Jury Innovations Improve Juror Understanding of DNA Evidence?*, 90 JUDICIATURE 152 (2007); David H. Kaye et. al., *Statistics in the Jury Box: How Jurors Respond to Mitochondrial DNA Match Probabilities*, 4 J. EMPIRICAL LEGAL STUD. 797-834 (2007); William C. Thompson, *Are Juries Competent to Evaluate Statistical Evidence?* 52 LAW & CONTEMP. PROB. 9 (1989); William C. Thompson & Edward Schumann, *Interpretation of Statistical Evidence in Criminal Trials: The Prosecutor's Fallacy and Defense Attorney's Fallacy*, 11 LAW AND HUM. BEHAV. 167 (1987).
- 196 See, e.g., *People v. Slone*, 76 Cal. App. 3d 611, 622 (Cal. Ct. App. 1978) (explicitly rejecting appellant's contention that the expert's conclusion that it was “highly probable” appellant's teeth created the bite mark at issue was scientifically invalid); *Bundy v. State*, 455 So. 2d 330, 349 (Fla. 1984) (“The evidence in question is based on the examination of impressions made by human teeth and their comparison with models of known human teeth for the purpose of determining whether the impressions were or probably were or could have been made by a particular individual ... the basis for the comparison testimony--that the science of odontology makes such comparison possible due to the significant uniqueness of individual dental characteristics--has been adequately established.”); see also *Kennedy v. State*, 640 P.2d 971, 978 (Okla. Crim. App. 1982) (“The means and techniques for making the models for comparison are complex, but they are based on standardized procedures known to produce accurate measurements.”).
- 197 397 N.W.2d 136 (Wis. Ct. App. 1986).

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE..., 4 Va. J. Crim. L. 1

- 198 369 S.E.2d 870 (W. Va. 1988).
- 199 See *Armstrong*, 369 S.E.2d at 874-77; *Stinson*, 397 N.W.2d at 139.
- 200 *Stinson*, 397 N.W.2d at 139-40 (holding that the “standards and controls” at issue were, of course, for the collection of the data, not the interpretation of the data collected).
- 201 See *supra* note 158-60. While we focus on these two hypotheses for purposes of this discussion, there are at least three other hypotheses underlying bite mark analysis: that forensic dentists are capable of distinguishing a bite mark from other pattern injuries; that the human dentition is unique; and, even assuming uniqueness, that human skin is capable of accurately recording the uniqueness. None of these hypotheses have been scientifically validated and, as discussed below, recent research tends to undermine them.
- 202 There is no **evidence** that the dentitions of twins are any more or less alike than any other adult dentitions.
- 203 *Stinson*, 397 N.W.2d at 137-39, 142 (emphasis added).
- 204 See *supra* note 17.
- 205 Craig A. Kubiak, Comment, *Scientific Evidence in Wisconsin: Using Reliability to Regulate Expert Testimony*, 74 MARQ. L. REV. 261, 279-80 (1991).
- 206 *Stinson*, 397 N.W.2d at 139 (emphasis added).
- 207 *Id.* at 140.
- 208 See N.J. Schweitzer & M.J. Saks, *The CSI Effect: Popular Fiction About Forensic Science Affects Public Expectations About Real Forensic Science*, 47 JURIMETRICS J. 357 (2007); Donald E. Shelton et al., *A Study of Juror Expectations and Demands Concerning Scientific Evidence: Does the ‘CSI Effect’ Exist?* 9 VAND. J. ENT. & TECH. L. 331 (2006); Max M. Houck, *CSI: The Reality*, SCI. AM., July 2006, at 85.
- 209 See William T. Carrington, *Mississippi Innocence: The Convictions and Exonerations of Levon Brooks and Kennedy Brewer and the Failure of the American Promise*, 28 G'TOWN J. LEGAL ETHICS 123 (2015).
- 210 *Stinson v. City of Milwaukee*, No. 09-C-1033, 2013 WL 5447916, at *12 (E.D. Wis. Sept. 30, 2013) (“*Stinson's* tooth 8, which was broken to the root, could not create a mark on the victim's skin without significant damage occurring.”).
- 211 Scholars have noted that bite mark examiners often fail to actually match bite marks to the dentition that made those bite marks, even in the context of controlled studies. Thus, as one forensic dentist has noted, bite mark **evidence** is subject to a “disturbingly high false-positive error rate.” Bowers, *supra* note 159, at S106. Such is **evidenced** by: a 1975 study which found that bite mark examiners made “incorrect identification[s] of ... bite[s]” on pig skin 24% of the time when the bites were made “under ideal laboratory conditions” and 91% of the time when “the bites were photographed 24 h[ours] after the bites were made;” a 1999 American Board of Forensic Odontology Bitemark Workshop “where ABFO diplomats attempted to match four bitemarks to seven dental models [and] found 63.5% false positives;” and, a 2001 study of “bites made in pig skin, widely accepted as an accurate analogue of human skin,” which resulted in 11.9-22.0% “false positive identifications ... for various groups of forensic odontologists.” *Id.*
- 212 See, e.g., NAS REPORT, *supra* note 4, at 176; H. David Sheets et al., *Dental Shape Match Rates in Selected and Orthodontically Treated Populations in New York State: A Two Dimensional Study*, 56 J. FORENSIC SCI. 621 (2011) [hereinafter Sheets et al., *Dental Shape Match Rates*].
- 213 See, e.g., *Handley v. State*, 515 So. 2d 121, 131 (Ala. Crim. App. 1987) (“Based upon our own precedent and the persuasiveness of other jurisdictions' rulings, we, too, hold that the admissibility of the dental witness's bite mark comparison does not depend on meeting the *Frye* standard. In the instant case, the jury itself was able to look at photographic overlays of the plastic

models of the bite marks and of appellant's teeth.”); *People v. Slone*, 76 Cal. App. 3d 611, 624 (Ct. App. 1978) (“The *Marx* court distinguished the bite mark evidentiary presentation from other scientific-test **evidence** ... on the ground that there was a more trustworthy basis for admissibility of the bite-mark-identification **evidence** ... due to the fact that the trier of fact could see for itself, by looking at the material-object exhibits ... what constituted the basis for comparison with a defendant's dentition.”); *Bundy v. State*, 455 So. 2d 330, 349 (Fla. 1984) (“With bite marks **evidence**, the jury is able to see the comparison for itself by looking directly at the physical **evidence** in the form of photographs and models ... The technique is similar to hair comparison **evidence**, which is admissible even though it does not result in identifications of absolute certainty as fingerprints do.”); *People v. Milone*, 356 N.E.2d 1350, 1358 (Ill. 1976) (“Another factor effecting the admissibility of scientific testimony involves the nature of the **evidence** being offered. In *Jennings*, the court refused to accept testimony based upon the workings of a machine (lie detector) which had not proved to be substantially reliable and the results of which were subject to various subjective interpretations. Bite mark comparison, on the other hand, involves only a visual comparison between the wound and the dentition of the defendant. The great care taken to preserve and gather the physical **evidence** in this case precludes any problems arising in regard to the quality of the exhibits being compared.

For this reason, the testimony of the experts serves only to lend assistance to the trial court in interpreting physical **evidence** not within the ken of the average trial judge's knowledge.”); *State v. Peoples*, 605 P.2d 135, 139 (Kan. 1980) (“The superior trustworthiness of the scientific bite mark approach ... is due to the fact that the trier of fact could see for itself ... what constituted the basis for comparison with a defendant's dentition.”); *Commonwealth v. Cifizzari*, 492 N.E.2d 357, 363, 363 n.15 (Mass. 1986) (“The admissibility of expert dental witnesses' testimony does not depend on meeting the *Frye* test. The experts' testimony merely aided the jury in comparing the photographs of the bite marks with the defendant's dental impressions We are not denigrating from *Frye* because we recognize the importance of establishing scientific reliability of new theories. We simply rule that *Frye* is not here applicable.”); *Kennedy v. State*, 640 P.2d 971, 977 (Okla. Crim. App. 1982) (“We cite with approval the leading California case on bite-mark identification [*Marx*]. There, the Court of Appeals emphasized that the bite-mark **evidence** was trustworthy because the basic data on which the experts based their conclusions were verifiable by the court. In *Marx*, as here, the trier of fact was shown models, photographs, and overlays of the victim's wounds and the accused teeth. The jury and the judge could see the extent to which the bite marks conformed to his teeth.”); *State v. Jones*, 259 S.E.2d 120, 124 (S.C. 1979) (internal quotation marks and citations omitted) (“In this case, we think admissibility depends upon ... the degree to which the trier of fact must accept, on faith, scientific hypotheses not capable of proof or disproof in a court and not even generally accepted outside the courtroom.”); *State v. Armstrong*, 369 S.E.2d 870, 876 (W. Va. 1988) (“Many of the courts have emphasized that the reliability of bite-mark **evidence**, unlike most scientific **evidence**, is, when presented properly in the particular case, readily apparent; it is a ‘common sense’ type of comparison of physical **evidence** which lends itself readily to verification and understanding. The judge and the jury can see the extent to which the bite mark conforms to the suspect's teeth.”).

214 509 U.S. 579 (1993). As discussed below, by the time the Supreme Court decided *Daubert*--and, six years later *Kumho Tire*--bitemark jurisprudence had been established; there exists not a single published decision applying *Daubert* analysis to bite mark **evidence**.

215 526 U.S. 137 (1999). It is important to note that the Supreme Court in *Kumho Tire* rejected the distinction between science and technical **evidence** for purposes of applying the *Daubert* test because such a distinction would be difficult to draw. The Court wrote that, “it would prove difficult, if not impossible, for judges to administer evidentiary rules under which a gatekeeping obligation depended upon a distinction between ‘scientific’ knowledge and ‘technical’ or ‘other specialized’ knowledge. There is no clear line that divides the one from the others.” *Id.* at 148.

216 *State v. Reid*, 757 A.2d 482, 487 (Conn. 2000); *Bundy*, 455 So. 2d at 349 (“The technique is similar to hair comparison **evidence**, which is admissible even though it does not result in identifications of absolute certainty as fingerprints do.”); *Kennedy*, 640 P.2d at 977 (“The jury and the judge could see the extent to which the bite marks conformed to [the defendant's] teeth.”).

217 *Marx* cited no precedent--and there appears to have been none-- for this “eyeball test” as it relates to bite mark **evidence**; yet, as discussed *infra*, it has also been used to admit hair comparison testimony. Instead, *Marx* cited *Frye* for the proposition that it applies only when “the trier of fact must accept, on faith, scientific hypotheses not capable of proof or disproof in court and not even generally accepted outside the courtroom.” *People v. Marx*, 54 Cal. App. 3d 100, 110 (Cal. Ct. App. 1975). The court reasoned that if there was no risk of overwhelming the trier of fact, then the court need not “sacrifice its independence

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE..., 4 Va. J. Crim. L. 1

in favor of deference to the expert.” *Id.* at 111. *Frye*, however, makes no such distinction. Rather, *Frye* applies “when the question involved does not lie within the range of common experience or common knowledge, but requires special experience or special knowledge.” *Frye v. United States*, 293 F. 1013, 1014 (D.C. 1923).

218 See *State v. Porter*, 698 A.2d 739, 746-48 (Conn. 1997).

219 *Id.* at 746.

220 *Id.* at 748.

221 *Id.*

222 *Id.*

223 *Id.*

224 *Id.*

225 *State v. West*, 877 A.2d 787, 805 (Conn. 2000) (internal quotation marks and citations omitted).

226 *State v. Reid*, 757 A.2d 482, 487 (Conn. 2000).

227 See *State v. Armstrong*, 369 S.E.2d 870, 875 (W. Va. 1988) (“This case presents a question of first impression for this Court, specifically, the admissibility of bite-mark **evidence**. All of the twenty-one jurisdictions which have specifically addressed this question in a reported opinion, where a qualified expert was involved, have held bitemark **evidence** to be admissible for positive identification purposes, and the general reliability of bite-mark comparison techniques has been sufficiently established, such that a hearing in each case to establish the general reliability thereof is not necessary. The courts have rejected challenges to bite-mark **evidence** based upon constitutional, evidentiary and scientific arguments.”).

228 See *supra* note 217.

229 *Frye v. United States*, 293 F. 1013, 1014 (D.C. Cir. 1923).

230 See, e.g., *People v. Slone*, 76 Cal. App. 3d 611, 624, 625 (Cal. Ct. App. 1978) (relying on testimony of three forensic odontologists which showed “bite-mark-identification technique had gained general acceptance in the scientific community of dentistry--the relevant scientific community involved”); *People v. Watson*, 75 Cal. App. 3d 384, 401-02 (Cal. Ct. App. 1977) (basing admission on *Marx*, and finding that, once a new scientific technique has been accepted by the court, it may not be disrupted without “**evidence** reflecting change in the attitude of the scientific community,” presumably forensic odontologists); *People v. Marx*, 54 Cal. App. 3d 100, 110 (Cal. Ct. App. 1975) (giving credence to the testimony of the state's experts, who were optimistic that dental identification techniques could be used to identify bite marks, implying that the relevant scientific community were the experts themselves); *People v. Smith*, 468 N.E.2d 879, 889 (N.Y. 1984) (basing admission on the claim that the technique of comparing one photo of a bite mark to another was sufficiently reliable and had been “accepted by the scientific community,” comprised of prosecution and defense experts who together “acknowledged the reliability and acceptance of photographic comparisons”); *People v. Middleton*, 429 N.E.2d 100, 103 (N.Y. 1981) (admitting **evidence** by finding that “the test is not whether a particular procedure is unanimously indorsed by the scientific community, but whether it is generally acceptable as reliable ... [;] [t]he techniques employed (photography, freezing of tissue specimens, the taking of dental molds, visual observation) are accepted and approved by the majority of the experts in the field”);

231 See *supra* note 230 and accompanying text.

232 See NAS REPORT, *supra* note 4, at iv-ix.

233 The NAS Report noted that skin is simply not a suitable medium to record bite marks: “bite marks on the skin will change over time and can be distorted by the elasticity of the skin, the unevenness of the surface bite, and swelling and healing.” *Id.* at 174. In addition, problems may also arise because of “distortions in photographs and changes over time in the dentition

of suspects.” *Id.* The report goes on to note that the first hypothesis underlying bite mark analysis (i.e., that an association can be made between a dentition and a bite mark) is flawed because--despite guidelines published by the American Board of Forensic Odontology, which provide for various methods of bite mark analysis--“[t]here is no science on the reproducibility of the different methods of analysis that lead to conclusions about the probability of a match.” *Id.* Indeed, “[e]ven when using the guidelines, different experts provide widely differing results and a high percentage of false positive matches of bite marks using controlled comparison studies.” *Id.* As to the second hypothesis--i.e., that a valid estimate of the probative value of a putative “match” can be stated--“there is no established science indicating what percentage of the population or subgroup of the population could also have produced the bite.” *Id.*

234 *See infra* notes 236-37.

235 *See id.*

236 *See* Mary A. Bush et al., *A Study of Multiple Bitemarks Inflicted in Human Skin by a Single Dentition Using Geometric Morphometric Analysis*, 211 J. FORENSIC SCI. INT’L 1 (2011) [hereinafter Bush et al., *Using Geometric Morphometric Analysis*]; Mary A. Bush, et al., *Biomechanical Factors in Human Dermal Bitemarks in a Cadaver Model*, 54 J. FORENSIC SCI. 167, 167 (2009) [hereinafter Bush et al., *Biomechanical Factors*]; Mary A. Bush et al., *Inquiry into the Scientific Basis For Bitemark Profiling and Arbitrary Distortion Compensation*, 55 J. FORENSIC SCI. 976, 976-83 (2010); Mary A. Bush et al., *Similarity and Match Rates of the Human Dentition In 3 Dimensions: Relevance to Bitemark Analysis*, 125 INT’L J. LEGAL MED. 779 (2011); Mary A. Bush et al., *Statistical Evidence for the Similarity of the Human Dentition*, 56 J. FORENSIC SCI. 118 (2011); Mary A. Bush et al., *The Response of Skin to Applied Stress: Investigation of Bitemark Distortion in a Cadaver Model*, 55 J. FORENSIC SCI. 71 (2010) [hereinafter Bush et al., *The Response of Skin to Applied Stress*]; Raymond G. Miller et al., *Uniqueness of the Dentition as Impressed in Human Skin: A Cadaver Model*, 54 J. FORENSIC SCI. 909 (2009); H. David Sheets et al., *Bitemarks: Distortion and Covariation of the Maxillary and Mandibular Dentition as Impressed in Human Skin*, 223 FORENSIC SCI. INT’L 202, 202-07 (2012); H. David Sheets et al., *Patterns of Variation and Match Rates of the Anterior Biting Dentition: Characteristics of a Database of 3D Scanned Dentitions*, 58 J. FORENSIC SCI. 60 (2013); Sheets et al., *Dental Shape Match Rates*, *supra* note 212; H. David Sheets & Mary A. Bush, *Mathematical Matching of a Dentition to Bitemarks: Use and Evaluation of Affine Methods*, 207 FORENSIC SCI. INT’L 111 (2011) [hereinafter Sheets & Bush, *Mathematical Matching*].

237 *See* Bush et al., *The Response of Skin to Applied Stress*, *supra* note 236, at 73 (finding no two bitemarks created by the same dentition were measurably identical; shorter teeth created indentations smaller than their actual width, some as much as 25% smaller); Sheets & Bush, *Mathematical Matching*, *supra* note 236 (finding that the matching of dentition to the bitemarks created was not possible within limits of repeatable measurements).

238 “Langer lines” is the term used to describe the direction within human skin along which the skin has the least flexibility. *See* Bush et al., *Biomechanical Factors*, *supra* note 236 (of the twenty-three bitemarks made for the experiment, none were visually or measurably identical); Bush et al., *Using Geometric Morphometric Analysis*, *supra* note 236 (a single dentition was used to create eighty-nine bitemarks, none of which matched the measurable shape of the dentition; the bitemarks were also compared to 411 other dentitions, showing the closest match to the bitemark was not always the teeth that created the mark); *see also* Iain A. Pretty, *Unresolved Issues in Bitemark Analysis*, in BITEMARK EVIDENCE: A COLOR ATLAS AND TEXT 547, 548 (Robert B.J. Dorian ed., 1st ed. 2005) (noting that “skin is a poor registration material”).

239 Sheets et al., *Dental Shape Match Rates*, *supra* note 212, at 621-26.

240 *Id.*

241 NAS REPORT, *supra* note 4, at 174.

242 *See supra* pp. 47-48.

243 *See id.* at 175 (“The effect of distortion on different comparison techniques is not fully understood and therefore has not been quantified.”).

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE..., 4 Va. J. Crim. L. 1

- 244 See *supra* notes 237-40 and accompanying text.
- 245 The ABFO has since retreated from its claim that bitemark experts can identify the unique source of an alleged bitemark, a development discussed more fully *infra*, at Part IV.A.
- 246 See Cary T. Oien, *Forensic Hair Comparison: Background Information for Interpretation*, 11 FORENSIC SCI. COMM. (2009).
- 247 See *supra* note 158-60. For example, ballistics, latent print analysis, and bite mark **evidence** rest on the same hypotheses, i.e., that a properly trained expert can make an association between a mark at a crime and provide a statistical valid expression of the probative value of such an association.
- 248 There is--and was--an abundance of **evidence** of the existence of error. In 2002, Bruce Budowle, the research director of the FBI DNA Laboratory, and Max Houck, an expert in hair microscopy and director of the forensics program at West Virginia University, published a study that reviewed human hair examinations within the FBI laboratory that underwent both microscopical comparison and mtDNA analysis between 1996 and 2000. See Max M. Houck & Bruce Budowle, *Correlation of Microscopic and Mitochondrial DNA Hair Comparisons*, 47 J. FORENSIC SCI. 964 (2002). Of the eighty cases in which FBI hair examiners found a positive microscopic association, nine cases resulted in exclusions when the same hairs were subjected to mitochondrial DNA testing. *Id.* at 964-66. According to the NAS Report, the study “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.” NAS REPORT, *supra* note 4, at 161.
- 249 See *Browning v. State*, 33 Miss. 47 (1857).
- 250 *Id.* at 56.
- 251 *Id.* at 58.
- 252 *Id.* at 84.
- 253 103 Mass. 412 (Mass. 1869).
- 254 *Id.* at 419.
- 255 *Id.* at 420.
- 256 *Id.*
- 257 *Id.*
- 258 See, e.g., Michael J. Saks, *Banishing Ipse Dixit: The Impact of Kumho Tire On Forensic Identification Science*, 57 WASH. & LEE L. REV. 879, 881 (2000). As noted by Professor Saks:
From the viewpoint of conventional science, the forensic identification sciences are contenders for being the shoddiest science offered to the courts. After being in business for nearly a century, they still have developed little that would be recognized as a scientific foundation and, consequently, have little basic science to apply to their operational activities. For much of the twentieth century, the courts readily admitted these fields, apparently because they were flying the banner of science and not because they presented sound data supporting their claims.
Id. at 879.
- 259 *Id.*
- 260 John I. Thornton & Joseph L. Peterson, *The General Assumptions and Rationale of Forensic Identification*, in SCIENCE IN THE LAW: FORENSIC SCIENCE ISSUES 35 (D.L. Faigman et. al. eds., 2002).

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE..., 4 Va. J. Crim. L. 1

- 261 The label itself is problematic, as it may accord to a discipline a level of gravitas that is undeserved. *See, e.g., United States v. Starzecpyzel*, 880 F. Supp. 1027, 1038 (S.D.N.Y. 1995) (noting that forensic document examination, despite its certification procedure and the like, cannot be regarded as scientific knowledge after *Daubert*).
- 262 This term has come under criticism recently because, among other things, the term itself is ambiguous and has different meanings for different analysts. *See, e.g., United States v. Taylor*, 663 F. Supp. 2d 1170 (D.N.M. 2009); *United State v. Glynn*, 578 F. Supp. 2d 567 (S.D.N.Y. 2008).
- 263 *See, e.g., State v. West*, 877 A.2d 787, 807 (Conn. 2005) (finding that hair comparison **evidence** is rooted in science but is not subject to a threshold reliability hearing because it “simply requires jurors to employ their own powers of observation and comparison”); *State v. Reid*, 757 A.2d 482, 487 (Conn. 2000) (finding that hair comparison **evidence** “is based in science” but is not subject to a threshold reliability hearing because it allows jurors “to make their own determinations as to the weight they would accord the expert's testimony in the light of [the **evidence**] ... and their own powers of observation and comparison”); *Murray v. State*, 3 So. 3d 1108, 1117 (Fla. 2009) (internal quotation marks and citations omitted) (reiterating that “[v]isual and microscopic hair comparison is not based on new or novel scientific principles and, therefore, does not require a Frye analysis”); *McDonald v. State*, 952 So. 2d 484, 498 (Fla. 2006) (finding visual and microscopic hair comparison does not require a *Frye* analysis because it “is not based on new or novel scientific principles”); *Jent v. State*, 408 So. 2d 1024 (Fla. 1981) (finding that hair analysis **evidence** is not so unreliable and scientifically unacceptable that it is error to admit it); *Beam v. State*, 463 S.E.2d 347, 349 (Ga. 1995) (finding that the crime lab expert's hair analysis is admissible because § 24-9-67 of the Official Code Of Georgia Annotated states that “the opinions of experts on any question of science ... or like questions shall always be admissible”); *People v. Harvey*, 568 N.E.2d 381, 387 (Ill. App. Ct. 1991) (finding the trial court, following a *voir dire* hearing to determine the number of scientific areas of consistency between the hair samples, acted within its discretion in admitting the **evidence**, despite the State's failure to establish that the hairs were identical to a mathematical certainty), *McGrew v. State*, 682 N.E.2d 1289 (Ind. 1997); *Johnson v. Commonwealth*, 12 S.W.3d 258 (Ky. 1999) (concluding that Kentucky trial courts may take judicial notice that hair comparison analysis is scientifically reliable); *Commonwealth v. Tarver*, 345 N.E.2d 671, 676-77 (Mass. 1975) (“It was sufficiently shown in the record that the use of microscopic examination has been generally accepted by the community of scientists involved.”); *People v. Browning*, 308 N.W.2d 264 (Mich. Ct. App. 1981); *People v. Collins*, 204 N.W.2d 290, 293-94 (Mich. Ct. App. 1972) (holding that the defense's objection to the State witness's opinion that there was a “reasonable scientific certainty” that hairs shared a common origin--an objection which was based on the fact that “inability to identify hair samples by microscopic analysis is universally recognized”--affected the weight rather than the admissibility of the **evidence**); *State v. Hudson*, 970 S.W.2d 855, 860 (Mo. Ct. App. 1998) (holding that plain error review was not appropriate where the court's determined that hair analysis was admissible after the defendant argued the hair analysis was inadmissible because the scientific principles were not generally accepted under *Frye*); *State v. Millisor*, No. 9-98-69, 1999 Ohio App. LEXIS 3542, at *12-13 (Ohio Ct. App. Aug. 4, 1999) (finding analyst's statement that hair samples are consistent to a reasonable scientific certainty admissible); *Williamson v. State*, 812 P.2d 384, 405 (Okla. Crim. App. 1991) (citing *Driskell v. State*, 659 P.2d 343, 356 (Okla. Crim. App. 1983) (stating that the court remained committed to its “position as expressed in *Driskell* ... which sanctioned the use of hair comparison **evidence** and the determination that any question about the procedures and conclusions drawn therefrom should be raised on cross-examination,” despite the fact that the appellant had asked the court to reconsider its acceptance of hair comparison analysis and had cited four studies in support of his position that the analysis “[did] not meet sufficient standards of scientific reliability”); *Commonwealth v. Chmiel*, 30 A.3d 1111, 1142 (Pa. 2011) (conceding that “a once-viable science [hair analysis] may lose its wide acceptance in the scientific community and may be challenged pursuant to Rule 702,” although the science was widely accepted at the time of the appellant's trial in 2002); *State v. Fagundes*, 614 P.2d 198 (Wash. Ct. App. 1980) (finding that it was in the discretion of the trial judge to admit hair analysis **evidence** following an analyst's testimony regarding her testing methods and their general acceptability in the scientific community); *State v. Hicks*, 549 N.W.2d 435, 437 (Wis. 1996) (acknowledging that an “[analyst] stated that ... to a reasonable degree of scientific certainty, the unknown Negro and Caucasian hair specimens ‘could have’ come from Hicks and D.F.”).
- 264 *See, e.g., State v. Fukusaku*, 946 P.2d 32, 44 (Haw. 1997) (affirming the trial court's refusal to apply *Daubert* scrutiny to hair trace **evidence** because of its overwhelming acceptance by criminal courts and noting that, because “the scientific principles and procedures underlying hair and fiber **evidence** are well-established and of proven reliability, the **evidence** in the present

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE..., 4 Va. J. Crim. L. 1

case [could] be treated as ‘technical knowledge’”); *McGrew*, 682 N.E.2d at 1292 (discussing how hair comparison **evidence** is not subject to *Daubert* scrutiny because the technique relies on observations made by persons with specialized knowledge rather than being a matter of scientific principles).

265 *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 146 (1999).

266 *See id.*

267 *See* N.J. Schweitzer & Michael J. Saks, *The Gatekeeper Effect: The Impact of Judges' Admissibility Decisions on the Persuasiveness of Expert Testimony*, 15 PSYCHOL. PUB. POL'Y & L. 1 (2009); *see also* *United States v. Addison*, 498 F.2d 741, 744 (D.C. Cir. 1974) (holding that expert scientific **evidence** may “assume a posture of mythic infallibility in the eyes of a jury of laymen”).

268 Indeed, this is reason that the National Academy of Sciences had the authority to issue its comprehensive assessment of the state of forensic “science.” As the NAS Report stated:

The law's greatest dilemma in its heavy reliance on forensic **evidence**, however, concerns the question of whether--and to what extent--there is science in any given “forensic science” discipline. Two very important questions should underlie the law's admission of and reliance upon forensic **evidence** in criminal trials: (1) the extent to which a particular forensic discipline is founded on a reliable scientific methodology that gives it the capacity to accurately analyze **evidence** and report findings and (2) the extent to which practitioners in a particular forensic discipline rely on human interpretation that could be tainted by error, the threat of bias, or the absence of sound operational procedures and robust performance standards. These questions are significant.

Thus, it matters a great deal whether an expert is qualified to testify about forensic **evidence** and whether the **evidence** is sufficiently reliable to merit a fact finder's reliance on the truth that it purports to support. Unfortunately, these important questions do not always produce satisfactory answers in judicial decisions pertaining to the admissibility of forensic science **evidence** proffered in criminal trials.

NAS REPORT, *supra* note 4, at 9.

269 The NIST research agenda discussed above should lead to developing a more scientific basis for these conclusions. That this basic research is only now being undertaken further demonstrates the inadequacy of the current state of scientific knowledge underlying these techniques.

270 An example of this critique as it relates to ballistics and toolmarks can be found in Itiel's Dror's piece, *How Can Francis Bacon Help Forensic Science? The Four Idols of Human Biases*:

The subjective and unspecified identification criterion of sufficient agreement is an example of *idola fori*. Furthermore, the AFTE Theory of Identification stipulation that the determination of ‘sufficient agreement is the product of the examiner's personal training, skills, and experience’ also involves *idola specus*--the subjective individual's experience determines decisions, rather than scientifically measurable criteria based on objective, quantifiable measurement divorced from and independent of the specific incidental individual who is making the observations.

Itiel E. Dror, *How Can Francis Bacon Help Forensic Science? The Four Idols of Human Biases*, 50 JURIMETRICS J. 93, 104 (2009) (emphasis removed).

271 *See* NAS REPORT, *supra* note 4, at 8 (“[A]lthough some techniques may be too imprecise to permit accurate identification of a specific individual, they may still provide useful and accurate information about questions of classification.”).

272 The authors are indebted to Professor Simon A. Cole for his insight into the role of science in validating the conclusions proffered by expert witnesses, which significantly advanced our thinking on this topic. *See also infra* Part III (discussing the limiting language of hair comparison associations).

273 *See, e.g., Knoll v. State*, 12 N.W. 369, 370 (Wis. 1882) (“The opinion of the witness as to the fact that the hair came from the head of the same person was not admissible on the ground that the inquiry related to a scientific subject--one which required peculiar knowledge or previous study and experience to give information about. ... The witness, then, could not testify to his opinion on the ground that the subject-matter of the inquiry related to a scientific subject, and was expert testimony.”).

- 274 Hair experts began incorporating statistics into their conclusions largely based on a single 1974 study by Barry Gaudette, a former hair examiner for the Royal Canadian Mounted Police. *See* B.D. Gaudette & E.S. Keeping, *An Attempt at Determining Probabilities in Human Scalp Hair Comparison*, 19 J. FORENSIC SCI. 599, 599-606 (1974). The use of this research to provide a statistical weight for hair examiners' conclusions has been entirely discredited. *See* NAS REPORT, *supra* note 4, at 23-24.
- 275 *See e.g.*, *United States v. Massey*, 594 F.2d 676, 679-81 (8th Cir. 1979) (reversing and remanding for new trial because, in light of the importance of the hair comparison **evidence**, the use of statistical **evidence** from the Canadian study by Gaudette and Keeping--which found there was a 1 in 4,500 possibility that the hair could have come from someone other than the defendant--was not harmless error); *Williamson v. Reynolds*, 904 F. Supp. 1529 (E.D. Okla. 1995) (finding reversible error when an expert cited Gaudette's studies--which estimated that the probability that two microscopically similar hairs came from two different sources was 1 in 4,500 and that the probability for this kind of error in pubic hair analysis was 1 in 800--because the implication was that the hairs belonged to the petitioner); *People v. Cooper*, 809 P.2d 865, 878-79 (Cal. 1991) ("Unlike fingerprint comparison, an absolute match is not possible when comparing hairs."); *Thompson v. State*, 539 A.2d 1052, 1057-59 (Del. 1988) (finding that hair comparison **evidence** does not create probable cause to arrest a suspect because it is universally acknowledged that hair comparison **evidence** is not a form of positive identification, though it may link a suspect to a crime); *Long v. State*, 689 So. 2d 1055, 1058 (Fla. 1997) (reversing defendant's conviction for insufficient **evidence** because "[h]air comparisons cannot constitute a basis for positive personal identification because hairs from two different people may have precisely the same characteristics"); *Jackson v. State*, 511 So. 2d 1047, 1049 (Fla. Dist. Ct. App. 1988) (reversing trial court's denial of defendant's motion for acquittal and vacating his conviction and sentence because the defendant's conviction hinged on hair comparison **evidence** which did not result in absolute identification); *People v. Linscott* 566 N.E.2d 1355, 1360 (Ill. 1991) (finding the state's use of Gaudette statistics and its expert's conclusively stating the hairs belonged to defendant constituted reversible error); *State v. Carlson* 267 N.W.2d 170, 176 (Minn. 1978) (holding that the Gaudette statistics were improperly received but were cumulative and unprejudicial); *State v. Scarlett*, 426 A.2d 25, 28 (N.H. 1978) (finding harmless error, despite the fact that information about the Gaudette study was erroneously admitted as double-hearsay--at least--when a witness cited the study for the proposition that there was a forty-five to one chance that consistent hair had different origins); *State v. Bridges*, 421 S.E.2d 806, 808 (N.C. App. 1992) (finding no reversible error when witness testified to two studies on the probability of matching caucasian hairs coming from two different sources; the court found that **evidence** was admissible but could not be used to positively identify a person.); *Crawford v. State*, 840 P.2d 627, 636 (Okla. Crim. App. 1992) (finding witness properly testified as to the limits of hair comparison analysis when she stated that she could not conclude that a hair belonged to a particular individual beyond a reasonable doubt.); *Brown v. State*, 751 P.2d 1078, 1080 (Okla. Crim. App. 1988) (finding harmless error in admitting the Gaudette statistics).
- 276 1942 was the year that the "Laboratory became an official FBI division." Kim Waggoner, *The FBI Laboratory: 75 Years of Forensic Science Service*, 9 FORENSIC SCI. COMM. (2007), available at https://www.fbi.gov/about-us/lab/forensic-science-communications/fscc/oct2007/research/2007_10_research01_test2.htm/2007_10_research01_test1.htm. The "official birthday" of the FBI Lab is November 24, 1932. *See* John F. Fox, Jr., *The Birth of the FBI's Technical Laboratory--1924 to 1935*, FED. BUREAU OF INVESTIGATION, <https://www.fbi.gov/about-us/history/highlights-of-history/articles/laboratory> (last visited Nov. 21, 2015).
- 277 *See supra* notes 263-67 and accompanying text.
- 278 *See supra* notes 274-75 and accompanying text.
- 279 *See, e.g.*, *United States v. Hickey*, 596 F.2d 1082, 1084 (1st. Cir. 1979) ("An FBI agent testified that some hairs found on one of the ski masks, sweater, and in the hair brush were 'microscopically identical' to the hairs of the defendant"); *Massey*, 594 F.2d at 678 ("Agent James Hilverda, an expert in microscopic analysis, testified that Carl Massey's hair was similar to three of the five hairs found in the blue ski mask in all categories of microscopic comparison."); *United States v. Holleman*, 575 F.2d 139, 145 (7th Cir. 1978) ("An expert from the FBI laboratory testified that he had examined the human hairs found on those items and compared them to Taylor's hair. They matched in every one of the twenty microscopic, identifiable characteristics."); *Pitts v. State*, 617 S.W.2d 849, 851 (Ark. 1981) ("The Negroid hair, when examined with a microscope, had 20 different characteristics. Sample specimens of Pitts's hair had exactly the same 20 characteristics."); *Padilla v. People*, 397 P.2d 741, 743

(Colo. 1964) (describing how witness matched the victim's hairs to those found in the defendant's car and stated "unequivocally that the hairs were from one and the same person."); *Beam v. State*, 463 S.E.2d 347, 348 (Ga. 1995) (discussing crime lab expert who found that hair recovered from a hat in a cab "microscopically matched" the defendant's hair); *Paxton v. State*, 282 S.E.2d 912, 916 (Ga. 1981) ("The expert found 15 matching characteristics in appellant's hair samples and those found on the victim's sheet and stocking."); *People v. Jones*, 528 N.E.2d 648, 653 (discussing expert who found that hairs in defendant's car and victims' hairs had 13 characteristics in common); *People v. Columbo* 455 N.E.2d 733, 791 (Ill. App. Ct. 1983) (criminalist testified that if he found hairs to be 99.9% similar and .1% dissimilar, he would classify the hairs as dissimilar; however, he testified that "[e]very portion of this hair matched up exactly"); *Waters v. State*, 234 A.2d 147 (Md. 1967) (expert found that hairs found in victim's home were "identical" to the defendant's); *People v. Watkins*, 259 N.W.2d 381, 384-85 (Mich. Ct. App. 1977) (expert witness testified that hair found on the victim's pants matched the defendant's in "fifteen points of comparison"); *State v. Farrow*, 386 A.2d 808, 815 (N.H. 1978) (expert witness could not positively identify the hair but found it was similar to defendant's "in all fifteen recognized microscopic characteristics"); *State v. Dickens*, 180 S.E.2d 844, 847 (N.C. 1971) (FBI agent testified that the hairs in question were "microscopically identical"); *State v. Barber*, 179 S.E.2d 404, 410 (N.C. 1971) (FBI agent testified hairs were "microscopically identical in all identifying characteristics"); *State v. Williams*, 657 S.W.2d 405, 410 (Tenn. 1983) (FBI agent testified that hairs taken from defendant's clothing and victim were "indistinguishable" and probably from a common source); *State v. Melson*, 638 S.W.2d 342, 349 (Tenn. 1982) (FBI agent testified that hairs from the victim's blouse "exactly matched" defendant's hair); *Ward v. State*, 427 S.W.2d 876, 880 (Tex. Crim. App. 1968) (expert witness testified that a pubic hair taken from the appellant was identical "in all characteristics" to those recovered from the victim's body); *State v. Golloday*, 470 P.2d 191, 205 (Wash. 1970) (state's expert testified that one of the hairs obtained from the victim was "microscopically identical" to defendant's pubic hair; defense expert testified that the hair was not even a pubic hair).

280 *See, e.g., Breen v. State*, 349 So. 2d 113, 117 (Ala. Crim. App. 1977) (noting that the witness, the Supervisor of Scientific Investigation for the Birmingham Police Department, "testified that there was a great deal of variation in hair diameter, color, length and texture on a single head and for this reason hair could not be specifically identified as belonging to a particular individual"); *People v. Allen*, 41 Cal. App. 3d 196, 202 (Cal. Ct. App. 1974) (criminologist "admitted on cross-examination that the present state of the art of testing hair presently made identification by hair samples inconclusive, as hair of any individual had a range of distinguishing characteristics"); NAT'L RESEARCH COUNCIL, DNA TECHNOLOGY IN FORENSIC SCIENCE 158 (1992) ("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 of opinion among equally qualified experts.").

281 Digital copies of these photographs were provided to the authors by Mr. Bromgard's attorney, Peter J. Neufeld. The original court exhibit remains in Mr. Neufeld's files.

282 Hsu, *Review of FBI Forensics Does Not Extend to Federally Trained State, Local Examiners*, *supra* note 79.

283 *See Jimmy Ray Bromgard*, INNOCENCE PROJECT, http://www.innocenceproject.org/Content/Jimmy_Ray_Bromgard.php (last visited Nov. 21, 2015).

284 *See id.*

285 *See id.*

286 *See supra* notes 274-75 and accompanying text.

287 Nevertheless, examples of the use of such statistics are as plentiful as they are invalid. *See, e.g., People v. Linscott*, 566 N.E.2d 1355, 1360 (Ill. 1991); *State v. Carlson*, 267 N.W.2d 170, 176 (Minn. 1978); *compare State v. Bromgard*, 862 P.2d 1140, 1141 (Mont. 1993) (expert testified that "in his experience the odds were one in one hundred that two people would have head hair or pubic hair so similar that they could not be distinguished by microscopic comparison and the odds of both head and pubic hair from two people being indistinguishable would be about one in ten thousand"), *with Jimmy Ray Bromgard*, *supra* note 252, and *State v. Bauer*, 683 P.2d 946, 951 (Mont. 1984) (upholding conviction because, among other independent **evidence**,

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE..., 4 Va. J. Crim. L. 1

analyst estimated that “the chances of another person having the same type of pubic and head hair were one in ten thousand”), and [Brown v. State](#), 751 P.2d 1078, 1080 (Okla. Crim. App. 1988) (introduction of Gaudette statistics found to be harmless).

288 *See supra* note 277 (citing examples of expert testimonial conclusions concerning the probative value of hair comparison **evidence**).

289 *Id.*

290 *See infra* Part III.

291 The FBI has labeled such testimony as a “Type 1 Error[.]” *See supra* note 120.

292 *See supra* notes 274-75.

293 This variety of false and misleading testimony has been labeled as a “Type 2 Error[.]” by the FBI. *See supra* note 120.

294 The FBI has labeled such testimony as a “Type 3 Error[.]” *See supra* note 120.

295 See the discussion of the Odom exoneration at Part I.C.

296 These errors have been classified as “Type 3” errors. *See supra* note 120.

297 *See supra* note 279 and accompanying text.

298 For examples of courts using “matching” interchangeably with “consistent with,” see [United States v. Barnes](#), 481 F. App'x 505 (11th Cir. 2012), [United States v. Reddicks](#), 237 F. App'x 826 (4th Cir. 2007); [Kogut v. Cnty. of Nassau](#), 894 F. Supp. 2d 230, 234 (E.D.N.Y. 2012); [Bessaha v. Rock](#), No. 09-CV-3581 (JFB), 2012 WL 1458195, at *1 (E.D.N.Y. 2012); [Williamson v. Reynolds](#), 904 F. Supp. 1529 (E.D. Okla. 1995), [Hinson v. State](#), 199 P.3d 1166 (Alaska Ct. App. 2008); [State v. Edgin](#), 520 P.2d 288 (Ariz. 1980); [People v. Tully](#), 282 P.3d 173 (Cal. 2012); [People v. Richardson](#), 183 P.3d 1146 (Cal. 2008); [People v. Higgins](#), 131 P.3d 995 (Cal. 2006); [People v. Pride](#), 833 P.2d 643 (Cal. 1992); [People v. Cooper](#), 809 P.2d 865 (Cal. 1991); [People v. Allen](#), 41 Cal. App. 3d 196 (Cal. Ct. App. 1974); [Waterhouse v. State](#), 82 So. 3d 84 (Fla. 2012); [State v. Fukusaku](#), 946 P.2d 32 (Haw. 1997); [State v. Liuafi](#), 623 P.2d 1271 (Haw. Ct. App. 1981); [People v. Hall](#), 194 Ill. 2d 305 (Ill. 2000); [People v. Moore](#), 171 Ill. 2d 74 (Ill. 2000); [People v. Harvey](#), 209 Ill. App. 3d 733 (Ill. App. Ct. 1988); [Brown v. State](#), 999 So. 2d 853 (Miss. Ct. App. 2008); [State v. Dexter](#), 954 S.W.2d 332 (Mo. 1997); [State v. Southern](#), 980 P.2d 3 (Mont. 1999); [State v. McNicholas](#), 369 S.E.2d 569 (N.C. 1988); [State v. Prevette](#), 345 S.E.2d 159 (N.C. 1988); [State v. Head](#), 338 S.E.2d 908 (N.C. 1986), [State v. Hannah](#), 322 S.E.2d 148 (N.C. 1984), [State v. Pratt](#), 295 S.E.2d 462 (N.C. 1982), [State v. Green](#), 290 S.E.2d 625 (N.C. 1982), [State v. Suddreth](#), 412 S.E.2d 126 (N.C. Ct. App. 1992), [State v. Faircloth](#), 394 S.E.2d 198 (N.C. Ct. App. 1990); [Bryan v. State](#), 935 P.2d 338 (Okla. Crim. App. 1997); [Harjo v. State](#), 882 P.2d 1067 (Okla. Crim. App. 1994); [Salazar v. State](#), 852 P.2d 729 (Okla. Crim. App. 1993); [Crawford v. State](#), 840 P.2d 627 (Okla. Crim. App. 1992); [Williamson v. State](#), 812 P.2d 384 (Okla. Crim. App. 1991); [Fox v. State](#), 779 P.2d 562 (Okla. Crim. App. 1989); [Brown v. State](#), 751 P.2d 1078 (Okla. Crim. App. 1988); [State v. Blair](#), No. E2008-00073-CCA-R3-CD, 2009 WL 4878615, at *1 (Tenn. Ct. App. 2009); [Arciba v. State](#), No. 10-08-00120-CR, 2009 WL 5155532, at *1 (Tex. App. 2009).

299 *See supra* note 298.

300 Transcript of Record at 1, 766, 823, 1033-34, [State v. Bridges](#), 90-CRS-23102 (Mecklenburg Cnty. Superior Ct. Jan. 22-29, 1991) (on file with authors) [hereinafter [Bridges Transcript](#)].

301 *Id.* at 773.

302 *Id.* at 839.

303 *Id.* at 803-04.

304 *Id.* at 825.

- 305 [State v. Bridges](#), 421 S.E.2d 806, 808 (N.C. Ct. App. 1992).
- 306 *Id.* at 809.
- 307 Bridges Transcript, *supra* note 300, at 999-1000.
- 308 While this article was in final revisions, Mr. Bridges's 1991 conviction was vacated, and he was released from prison after 25 years. The court found that introduction of the hair comparison in the case violated Mr. Bridges's due process rights under both the federal and North Carolina constitutions. *See* Consent Order, *State v. Bridges*, No. 90-CRS-23102-04 (Mecklenburg Cnty. Superior Ct. Oct. 1, 2015) (“The admission of testimony containing the identified error types [identified in the FBI review] violated the Defendant's right to due process because it exceeded the limits of science and overstated the significance of the hair analysis to the jury.”) (on file with authors); Order Directing Defendant's Release, *State v. Bridges*, No. 90-CRS-23102-04 (Mecklenburg Cnty. Superior Ct. Oct. 1, 2015) (on file with authors); Notice of Dismissal, No. 90-CRS-23102-04 (Mecklenburg Cnty. Superior Ct. Feb. 16, 2016)
- 309 *See* 51% of 300 DNA Exonerations Involved Use of Improper/Unvalidated Forensic Science: Breakdown by Discipline, INNOCENCE PROJECT, <http://www.innocenceproject.org/causes-wrongful-onviction/FSBreakdownDiscipline.pdf> (last visited Nov. 21, 2015).
- 310 *See, e.g.*, Garrett & Neufeld, *supra* note 10.
- 311 *See, e.g.*, [Johnson v. Commonwealth](#), 12 S.W.3d 258, 262-63 (Ky. 1999) (“Based upon the overwhelming acceptance of this **evidence** by other jurisdictions, as well as our own history of routine admission of this **evidence** at trial, trial courts in Kentucky can take judicial notice that this particular method or technique is deemed scientifically reliable.”).
- 312 757 A.2d 482 (Conn. 2000).
- 313 *See* [State v. Porter](#), 698 A.2d 739, 746 (Conn. 1997) (adopting the *Daubert* test in Connecticut).
- 314 Although such hearings are referred to as “*Porter* hearings” in Connecticut, we use the label “*Daubert* hearing” in states that have adopted the *Daubert* test for consistency.
- 315 [Reid](#), 757 A.2d at 487-88.
- 316 *Id.* at 487. In an apparent effort to foreclose any future challenge to the admissibility of hair comparison **evidence**, the court found that, “even if a [*Daubert/Porter*] hearing were necessary, ... microscopic hair analysis satisfied the *Porter* test because of its general acceptance in the scientific community.” *Id.* at 488 n.3. The court further found that: in *Kumho* ... the United States Supreme Court held that a trial court has discretion to apply *Daubert* to all expert testimony, not just that which constitutes ‘scientific **evidence**.’ We need not decide in this case whether to apply *Kumho* in our *Porter* analysis, however, because it would not alter our conclusion that the trial court properly admitted the **evidence**. *Id.* at 488 n.4.
- 317 *See* [Reid v. State](#), No. CV020818851, 2003 WL 21235422, at *1 (Conn. May 14, 2003) (granting defendant's petition for new trial); *see also* [Mark Reid](#), NAT'L REGISTRY OF EXONERATIONS, <https://www.law.umich.edu/special/exoneration/Pages/casedetail.aspx?caseid=3569> (last visited Oct. 9, 2015).
- 318 *See* [Reid](#), 2003 WL 21235422, at *7. The victim had been forcibly raped in a dark, wooded area around 1:00 A.M. and described the perpetrator as a black man between 5# 6# and 5#7# with freckles across his nose and under his eyes. *See id.* at *4. Mr. Reid was six feet tall and had no freckles. *See id.* at *7. The victim had been drinking prior to the incident. *See id.* at *3. Based on the location of the incident and the description of the perpetrator, the police suspected Mr. Reid. *See id.* at *5-6. Reid's photo was placed in a photo array, and the victim ultimately picked out his photo; the victim likewise identified him in court. *See id.* at *4.
- 319 *See id.* at *3-4.

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE..., 4 Va. J. Crim. L. 1

- 320 *Id.* at *5.
- 321 *Id.*
- 322 *Id.* at *12.
- 323 See *State v. Reid*, 757 A.2d 482, 487 (Conn. 2000).
- 324 See *id.* at 487; *Williamson v. Ward*, 110 F.3d 1508 (10th Cir. 1997).
- 325 Williamson and his co-defendant Dennis Fritz were both innocent. See *Ron Williamson*, INNOCENCE PROJECT, http://www.innocenceproject.org/Content/Ron_Williamson.php (last visited Nov. 21, 2015).
- 326 *State v. Reid*, 757 A.2d 482, 487 (Conn. 2000).
- 327 The district court was “unsuccessful in its attempts to locate any indication that expert hair comparison testimony meets any of the [*Daubert*] requirements.” *Williamson v. Reynolds*, 904 F. Supp. 1529, 1554 (E.D. Okla. 1995). The court further observed that “[a]lthough the hair expert may have followed procedures accepted in the community of hair experts, the human hair comparison results in this case were, nonetheless, scientifically unreliable.” *Id.* at 1554.
- 328 See *Williamson v. Ward*, 110 F.3d 1508, 1523 (10th Cir. 1997).
- 329 See *id.*
- 330 See generally JOHN GRISHAM, *THE INNOCENT MAN* (2006). DNA testing revealed that none of the hairs that hair microscopy experts had labeled “matches” belonged to the defendants. See *Ron Williamson*, *supra* note 325. In addition, a DNA profile developed from the semen evidence matched a third person, who had been one of the state's witnesses at trial. See *id.* Ron Williamson--and Dennis Fritz, who was also charged and convicted--were exonerated and released in April 1999. See *id.* At one point, Williamson was within five days of execution. *Id.* Collectively, the two spent eleven years imprisoned. See *id.*
- 331 See *McGrew v. State*, 673 N.E.2d 787, 800 (Ind. Ct. App. 1996) (“Mindful of our Supreme Court's endorsement of *Daubert* and its progeny, we nonetheless conclude that *Williamson* is inapplicable to this case.”); *Bryon v. State*, 935 P.2d 338, 359 n.62 (Okla. Crim. App. 1997) (“*Williamson* is not binding on this Court. Bryon offers no other reason to review this settled area of law.”); *State v. Fukusaku*, 946 P.2d 32, 43 n.5 (Haw. 1997).
- 332 See *Bryon*, 935 P.2d at 359 n.62.
- 333 *Calvin Lee Scott*, INNOCENCE PROJECT, http://www.innocenceproject.org/Content/Calvin_Lee_Scott.php (last visited Nov. 21, 2015); *Curtis McCarty*, INNOCENCE PROJECT, http://www.innocenceproject.org/Content/Curtis_McCarty.php (last visited Nov. 21, 2015); *David Johns Bryson*, INNOCENCE PROJECT, http://www.innocenceproject.org/Content/David_Johns_Bryson.php (last visited Nov. 21, 2015); *Dennis Fritz*, INNOCENCE PROJECT, http://www.innocenceproject.org/Content/Dennis_Fritz.php (last visited Nov. 21, 2015); *Jeffery Todd Pierce*, INNOCENCE PROJECT, http://www.innocenceproject.org/Content/Jeffrey_Pierce.php (last visited Nov. 21, 2015); *Robert Lee Miller, Jr.*, NAT'L REGISTRY OF EXONERATIONS, <http://www.law.umich.edu/special/exoneration/Pages/casedetail.aspx?caseid=3473> (last visited Nov. 21, 2015); *Ronald Keith Williamson*, NAT'L REGISTRY OF EXONERATIONS, <https://www.law.umich.edu/special/exoneration/Pages/casedetail.aspx?caseid=3752> (last visited Nov. 21, 2015); *Sedrick Courtney*, INNOCENCE PROJECT, http://www.innocenceproject.org/Content/Sedrick_Courtney.php (last visited Nov. 21, 2015).
- 334 See *McCarty v. State*, 904 P.2d 110, 125 (Okla. Crim. App. 1995) (“[McCarty] acknowledges that hair comparison evidence is routinely used in criminal trials and this Court has previously found such testimony to be admissible However, he urges this Court to reconsider its position regarding the admissibility of hair analysis evidence, a request rejected by this Court in the past. [McCarty] has not persuaded this Court to now hold otherwise.”) (internal citations omitted).
- 335 See *Curtis McCarty*, *supra* note 333.

- 336 *McCarty*, 904 P.2d at 125.
- 337 *See id.*
- 338 *State v. Coolidge*, 260 A.2d 547, 560 (N.H. 1969) (quoting *Frye v. United States*, 293 F. 1013, 1014 (D.C. Cir. 1923)).
- 339 *See State v. Scarlett*, 426 A.2d 25, 27-28 (N.H. 1981) (relying on *Farrow* and distinguishing *Coolidge*, the court determined that expert testimony that (1) hair found on defendant's bed and the victim's hair were "morphologically similar" in "fifteen recognized microscopic characteristics," and (2) that when hairs are "found to be consistent with respect to all these different microscopic characteristics ... the chances of them having come from anyone else are forty-five to one" was admissible); *State v. Farrow*, 386 A.2d 808, 815 (N.H. 1978) (relying on *Breest*, the court admitted evidence where the expert witness "could not positively identify that the hair" belonged to defendant, but could conclude that the hair was similar to the defendant's "in all fifteen recognized microscopic characteristics"); *State v. Breest*, 367 A.2d 1320, 1331-33 (N.H. 1976) (rejecting a due process challenge to an expert witness' hair comparison and identification testimony that there "exists a high degree of probability and reasonable ability that we have had contact between this (victim's) clothing and that (defendant's) car").
- 340 *State v. Roff*, 67 A.D.2d 805, 805 (N.Y. 1979).
- 341 *See, e.g.,* *People v. Allweiss*, 396 N.E.2d 736 (N.Y. 1979); *Matter of Barber v. Rubin*, 72 A.D.2d 347, 350 (N.Y. App. Div. 1980) ("[A]n expert in the field can conclude with a reasonable degree of certainty whether hair from an unknown source matches the hair from a known source; that hair samples microscopically alike or closely similar can be said with a high degree of probability to have originated from the same source.").
- 342 *See* Spencer S. Hsu, *Review of FBI Forensics Does Not Extend to Federally Trained State, Local Examiners*, *supra* note 79.
- 343 *See* Bridges Transcript, *supra* note 300, at 766.
- 344 *See* Hsu, *Review of FBI Forensics Does Not Extend to Federally Trained State, Local Examiners*, *supra* note 79.
- 345 *Id.*
- 346 *Id.*
- 347 *See* Symposium, *Proceedings on the International Symposium on Forensic Hair Comparisons*, DEP'T OF JUSTICE (June 25-27, 1985), available at <https://www.ncjrs.gov/pdffiles1/Digitization/116592NCJRS.pdf>.
- 348 *See id.* at 193.
- 349 *Id.* at 199.
- 350 *Id.*
- 351 *Id.* at 204.
- 352 *Id.* at 209.
- 353 *Id.*
- 354 This deposition occurred on January 10, 2013 in *Gates v. Dist. of Columbia*, No. 1:11-cv-00040, at 330 (D.D.C) (on file with authors), which involves a claim of wrongful conviction based on hair comparison testimony. *See* Spencer S. Hsu, *D.C. Police Framed Man Imprisoned 27 Years for 1981 Murder, U.S. Jury Finds*, WASH. POST (Nov. 18, 2015), https://www.washingtonpost.com/local/public-safety/dc-police-framed-man-imprisoned-27-years-for-1981-murder-us-jury-finds/2015/11/18/7a454054-8e04-11e5-baf4-bdf37355da0c_story.html. The plaintiff was convicted of rape and murder, in part based on hair comparison testimony by an FBI examiner, and spent over twenty years in prison for those crimes. *See id.* Mr. Gates ultimately prevailed in his wrongful conviction suit in November of 2015. *See id.*

- 355 *See* Oien, *supra* note 246.
- 356 *Id.*
- 357 *See* NAS REPORT, *supra* note 4.
- 358 *Id.* at 156.
- 359 *Id.* at 160.
- 360 *Id.*
- 361 *Id.* at 161.
- 362 Hsu, *Convicted Defendants Left Uninformed of Forensic Flaws Found by Justice Department*, *supra* note 76.
- 363 *Id.*
- 364 *Id.*
- 365 *See supra* Part III.B; *see also* Yamil Berard, *Forensic Science Commission to Review Convictions Based on Hair Samples*, FORENSIC (Aug. 13, 2013), <http://www.forensicmag.com/news/2013/08/forensic-science-commission-reviewconvictions-based-hair-samples>.
- 366 *See* Norman L. Reimer, *The Hair Microscopy Review Project: An Historic Breakthrough for Law Enforcement and a Daunting Challenge for the Defense Bar*, NAT'L ASS'N OF CRIM. DEFENSE LAWYERS, (July 2013), <https://www.nacdl.org/champion.aspx?id=29488>. The review was announced in July 2013 and is being administered in conjunction with the Innocence Project and the National Association of Criminal Defense Lawyers. *See also* Press Release, Innocence Project, Innocence Project and NACDL Announce Historic Partnership with the FBI and Department of Justice on Microscopic Hair Analysis Cases, (July 18, 2013), *available at* <http://www.innocenceproject.org/news-events-exonerations/press-releases/innocence-project-and-nacdl-announce-historic-partnership-with-the-fbi-and-department-of-justiceon-microscopic-hair-analysis-cases>.
- 367 FED. BUREAU OF INVESTIGATION, MICROSCOPIC HAIR COMPARISON ANALYSIS AGREEMENT, *supra* note 120, at 1.
- 368 *Id.*
- 369 *See id.*
- 370 *See id.* (discussing error types).
- 371 *See* Reimer, *supra* note 366.
- 372 *See infra* note 428.
- 373 *See* Robert Aronson & Jacqueline McMurtrie, *The Use and Misuse of High-Tech Evidence by Prosecutors: Ethical and Evidentiary Issues*, 76 FORDHAM L. REV. 1453, 1487 (2007).
- 374 AM. JUDICATURE SOC'Y, CONFERENCE ON PREVENTING THE CONVICTION OF INNOCENT PERSONS 5 (2003).
- 375 Indeed, hair microscopy and bite mark analysis still enjoy near universal admissibility. *See supra* Parts II-III.
- 376 For a good overview of several incidents of systemic forensic fraud, *see* Paul C. Giannelli, *Scientific Fraud*, 46 CRIM. L. BULL. 1313 (2010).

- 377 See, e.g., Bridges Transcript, *supra* note 300, at 999-1000.
- 378 See, Paul C. Giannelli & Kevin C. McMunigal, *Prosecutors, Ethics, and Expert Witnesses*, 76 *FORDHAM L. REV.* 1493 (2007). There are several documented incidents, each of which also reflects an acknowledgement of the damage to the system as a whole. The FBI, for example, which runs the nation's most renowned forensic lab, had one of its own analysts plead guilty to a charge of making false statements, an occurrence which, according to the Justice Department's Inspector General, "has damaged intangibly the credibility of the FBI laboratory" as a whole. Maurice Possley et al., *Scandal Touches Even Elite Labs*, *CHI. TRIBUNE* (Oct. 21, 2004), <http://www.chicagotribune.com/news/watchdog/chi-041021forensics-story.html>. In addition, some courts have authorized administrative actions into incidences of gross forensic malfeasance. See *In re Investigation of the W. Va. State Police Crime Lab., Serology Div.*, 438 S.E.2d 501 (W. Va. 1993) (discussing the fact that the West Virginia Supreme Court appointed a special judge to investigate claims that Fred Zain, a serologist in the West Virginia State Police Forensic Laboratory, gave false testimony and sometimes offered comment about the effect of gross forensic malfeasance). Courts have also on occasion--often in dissent--offered criticism. See, e.g., *Brooks v. State*, 748 So. 2d 736, 750 (Miss. 2006) (McRae, J., dissenting) (internal citations omitted) ("This Court's apparent willingness to allow West to testify to anything and everything so long as the defense is permitted to cross-examine him may be expedient for prosecutors but it is harmful to the criminal justice system.").
- 379 Leaving aside the issue that arises from the admission of false **evidence**, there is a good deal of scholarship that specifically addresses the effectiveness of cross-examination regarding forensic **evidence**. See, e.g., Jonathan Koehler, *If the Shoe Fits They Might Acquit: The Value of Forensic Science Testimony*, 8 *J. EMPIRICAL STUD.* 21 (2011). Indeed, one experiment, which tested the reaction of potential jurors to flaws in microscopic hair examination, found that alerting jurors to problems had little impact on their decision-making. See Dawn McQuiston-Surrett and Michael J. Saks, *Communicating Opinion Evidence in the Forensic Identification Sciences: Accuracy and Impact*, 59 *HASTINGS L.J.* 1159, 1167-69 (2008) ("Whether or not jurors were informed about the limitations of microscopic hair examination on cross-examination or by the judge had little measurable or meaningful impact on their judgments about the likelihood that the defendant was the source of the crime-scene hair or their perceived understanding of the expert's testimony.").
- 380 For further support of this proposition, see *Hinton v. Alabama*, 134 S. Ct. 1081, 1084 (2014) (holding that where the only **evidence** linking defendant to crime was ballistic **evidence**, counsel was ineffective for failing to seek additional funds, which he wrongly believed were not available, to hire a better qualified expert to rebut prosecution's expert). The following is a non-exhaustive list of authorities finding counsel ineffective for various failures in responding to scientific **evidence** or expert testimony: *Williams v. Thaler*, 684 F.3d 597, 604 (5th Cir. 2012) (holding that trial counsel was ineffective in his failure "to obtain any independent ballistics or forensics experts, and was therefore unable to offer any meaningful challenge to the findings and conclusions of the state's experts, many of which proved to be incorrect"); *Elmore v. Ozmint*, 661 F.3d 783, 851, 864, 869-72 (4th Cir. 2011) (holding that failure of petitioner's lawyers to investigate state's forensic **evidence**--including hair and serology **evidence**--constituted ineffective assistance of counsel); *Duncan v. Ornoski*, 528 F.3d 1222, 1235 (9th Cir. 2008) ("[W]hen the prosecutor's expert witness testifies about pivotal **evidence** or directly contradicts the defense theory, defense counsel's failure to present expert testimony on that matter may constitute deficient performance."); *Bell v. Miller*, 500 F.3d 149, 155-57 (2d Cir. 2007) (holding that counsel was ineffective for failing to consult medical expert regarding reliability of shooting victim's identification); *Richey v. Bradshaw*, 498 F.3d 344, 361-64 (6th Cir. 2007) (holding that defendant's counsel provided ineffective assistance in **arson** trial where counsel failed to properly cross-examine the state's experts or to present competing scientific **evidence**); *Ege v. Yukins*, 485 F.3d 364, 378-80 (6th Cir. 2007) (holding that counsel's failure to object to admission of bite mark testimony constituted deficient performance--as required for petitioner to meet the "cause" prong of the cause-and-prejudice standard for review of the defaulted habeas claim--where bite mark was only physical **evidence** connecting her to crime scene); *Dugas v. Coplan*, 428 F.3d 317, 323, 328-31, 341-42 (1st Cir. 2005) (holding that counsel was ineffective in failing to pursue a "not **arson**" defense where state's strongest **evidence** was expert testimony on **arson**); *Draughon v. Dretke*, 427 F.3d 286, 296-97 (5th Cir. 2005) (holding that counsel was ineffective in failing to offer expert ballistics **evidence** in defense where government's prosecution was based on testimony from ballistics expert); *Soffar v. Dretke*, 368 F.3d 441, 443, 473-78 (5th Cir. 2004) (holding that defense counsel's failure to pursue and develop expert ballistics testimony that would have presented the jury with conflicting **evidence** bearing on the defendant's role in the crime was ineffective); *United States v. Tarricone*, 996 F.2d 1414, 1418-20 (2nd Cir. 1993) (holding that defense attorney was ineffective

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE..., 4 Va. J. Crim. L. 1

for failing to consult a handwriting expert who might have proven that the defendant never signed the agreement at issue in case);

Sims v. Livesay, 970 F.2d 1575, 1580-81 (6th Cir. 1992) (holding that counsel was ineffective for not hiring expert to analyze the bullet holes and powder patterns on a quilt the homicide victim held in her hands at the time of her shooting); *Sturgeon v. Quarterman*, 615 F. Supp. 2d 546, 572 (S.D. Tex. 2009) (finding ineffective assistance of counsel in failure to prepare expert witness to testify about unreliability of eyewitness identification); *State v. Smith*, 85 So. 3d 1063, 1083 (Ala. Crim. App. 2010) (finding ineffectiveness supported by lack of expert testimony relating to police procedures); *State v. Fitzpatrick*, 118 So. 3d 737, 753-57, 759, 762-64, 769-70 (Fla. 2013) (holding that counsel was ineffective for failing to adequately investigate and obtain expert assistance to rebut state's forensic expert's testimony); *Commonwealth v. Bussell*, 226 S.W.3d 96, 105 (Ky. 2007) (holding that counsel was ineffective for failing to adequately investigate and obtain expert assistance to rebut state's forensic expert testimony); *Wolfe v. State*, 96 S.W.3d 90, 94-95 (Mo. 2003) (finding that counsel was ineffective for failing to test hair samples found in victim's car); *Cravens v. State*, 50 S.W.3d 290, 295 (Mo. Ct. App. 2001) (holding that counsel was ineffective in failing to locate and present expert witnesses on forensic pathology and bullet analysis); *Wilhoit v. State*, 816 P.2d 545, 546-47 (Okla. 1991) (holding that counsel's failure to investigate bite-mark **evidence** constituted ineffective assistance of counsel); *Ard v. Catoe*, 372 S.C. 318, 327, 330-31, 336 (S.C. 2007) (finding that counsel was ineffective for failing to adequately develop and present gunshot residue **evidence** in response to government's expert testimony); *but see United States v. Higgs*, 663 F.3d 726, 738 (4th Cir. 2011) (finding no *Strickland* violation where "counsel conducted a thorough and effective cross-examination ... demonstrating that [he was] well acquainted with the criticisms" of the forensic discipline at issue); *United States v. Davis*, 406 F.3d 505, 509 (8th Cir. 2005) ("Davis's trial counsel cannot be said to be ineffective for failing to challenge the FBI's methodology on a basis that was not advanced by the scientific community at the time of trial."); *Libby v. McDaniel*, No. 3:04-CV-0038-LRH-RAM, 2011 WL 1301537, at *1, 9 (D. Nev. Mar. 31, 2011) (finding no ineffective assistance of counsel where defendant "offered no **evidence** that, as of 1990, the research or expertise necessary to successfully challenge forensic **evidence** was reasonably available to trial counsel"); *Wyatt v. State*, 71 So. 3d 86, 103 (Fla. 2011) (finding no ineffective assistance of counsel where the flaws in the forensic discipline were not known until well after defendant's trial); *Robertson v. State*, No. M2007-01378-CCA-R3-PC, 2009 WL 277073, at *1, 17 (Tenn. Crim. App. Feb. 5, 2009) (finding no ineffective assistance of counsel where, at time of trial, counsel "did not have the benefit of the FBI's retraction").

381 *See Cornell v. 360 W. 51st St. Realty*, 22 N.Y.3d 762, 785-86 (N.Y. 2014) ("[S]cientific understanding, unlike a trial record, is not by its nature static; the scientific consensus prevailing at the time of the Frye hearing in a particular case may or may not endure."); *see also supra* Parts II-III (discussing cases that took judicial notice of hair and bite mark **evidence**).

382 Our argument is in accord with the American Society of Crime Lab Directors/Laboratory Accreditation Board, which issued the following statement in response to the joint FBI/DOJ hair microscopy case audit:
We have an ethical obligation to take appropriate action if there is potential for, or there has been, a miscarriage of justice due to circumstances that have come to light, incompetent practice or malpractice. It is not ASCLD/LAB's intent to direct that such reviews be conducted by any laboratory or judicial system but it is our recommendation that each laboratory, in consultation with the appropriate legal authorities, consider whether there may be past cases, specifically involving convictions, in which it would be appropriate to evaluate the potential impact of the reported conclusions and/or related testimony on the conviction. Press Release, Am. Soc'y of Crime Lab. Directors/Lab. Accreditation Board, Notification from the ASCLD/LAB Board of Directors to Interested Parties Concerning Potential Issues with Hair Comparison Testimony (Apr. 21, 2013), *available at* <http://www.ascl-dlab.org/notification-from-the-ascl-dlab-board-of-directors-to-interested-parties-concerning-potential-issues-with-hair-comparison-testimony/> [hereinafter ASCLD Press Release].

383 *See, e.g.,* AM. BAR. ASS'N, REPORT TO THE HOUSE OF DELEGATES I (2004) ("[Counsel should] have competence in the relevant area or consult with those who do where forensic **evidence** is essential in a case."); *see generally* AM. BAR ASS'N, ACHIEVING JUSTICE: FREEING THE INNOCENT, CONVICTING THE GUILTY (2006) [hereinafter ACHIEVING JUSTICE].

384 *See* ACHIEVING JUSTICE, *supra* note 383378, at 99; Giannelli & McMunigal, *supra* note 344; Aronson & McMurtrie, *supra* note 373.

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE..., 4 Va. J. Crim. L. 1

- 385 See INNOCENCE PROJECT & WINSTON & STRAWN LLP, MEMORANDUM OF POTENTIAL POST-CONVICTION ARGUMENTS AND AUTHORITY BASED ON DISCREDITED HAIR MICROSCOPY ANALYSIS (2014), available at http://www.americanbar.org/content/dam/aba/events/criminal_justice/Forensics_Update_Post_Conviction_Discredited_Science.authcheckdam.pdf [hereinafter Post-Conviction Memorandum].
- 386 See Giannelli & McMunigal, *supra* note 378, at 1501-06; see also *N. Mariana Islands v. Bowie*, 243 F.3d 1109, 1118 (9th Cir. 2001) (“[A prosecutor’s due process duty] requires a prosecutor to act when put on notice of the real possibility of false testimony. This duty is not discharged by attempting to finesse the problem by pressing ahead without a diligent and a good faith attempt to resolve it. A prosecutor cannot avoid this obligation by refusing to search for the truth and remaining willfully ignorant of the facts.”).
- 387 MODEL RULES OF PROF’L CONDUCT R. 3.3 (1983). Note, however, that Comment 8 to Rule 3.3 clarifies the parameters of knowing presentation of false evidence by stating that, “[t]he prohibition against offering false evidence only applies if the lawyer knows that the evidence is false. A lawyer’s reasonable belief that evidence is false does not preclude its presentation to the trier of fact.” MODEL RULES OF PROF’L CONDUCT R. 3.3 cmt. 8.
- 388 MODEL RULES OF PROF’L CONDUCT R. 3.3.
- 389 MODEL RULES OF PROF’L CONDUCT R. 3.3 cmt. 13.
- 390 *Id.*
- 391 There is no prohibition against offering such measures anyway, regardless of the passing of the time frame. See MODEL RULES OF PROF’L CONDUCT R. 3.3. The authors have not seen a case where any lawyer has taken such steps in these circumstances.
- 392 See e.g., David S. Caudill, *Lawyers Judging Experts: Oversimplifying Science and Undervaluing Advocacy to Construct an Ethical Duty?*, 38 PEPP. L. REV. 675 (2011); Giannelli & McMunigal, *supra* note 378, at 1535.
- 393 MODEL RULES OF PROF’L CONDUCT R. 1.1 (1983). The rule requires “competent representation to a client,” which is defined as “the legal knowledge, skill, thoroughness and preparation reasonably necessary for the representation.” *Id.*
- 394 See, e.g., KATHLEEN M. RIDOLFI & MAURICE POSSLEY, N. CAL. INNOCENCE PROJECT, PREVENTABLE ERROR: A REPORT ON PROSECUTORIAL MISCONDUCT IN CALIFORNIA 1997-2009 (2010).
- 395 See, e.g., *Imbler v. Pachtman*, 424 U.S. 409 (1976).
- 396 See *Connick v. Thompson*, 563 U.S. 51 (2011) (holding that there can be no municipal liability for a district attorney’s office’s failure to train its prosecutors to turn over exculpatory (*Brady*) evidence on the basis of a single violation of that obligation); *Van de Kamp v. Goldstein*, 555 U.S. 335, 344 (2009) (holding that extended immunity included the concededly administrative acts of a district attorney’s office’s supervisory prosecutors in systemic “training or supervision or information-system management”); see also Ted Sampsell-Jones & Jenna Yauch, *Official and Municipal Liability for Constitutional and International Torts Today: Does the Roberts Court Have An Agenda?*, 80 FORDHAM L. REV. 623 (2011). See also Barry Scheck, *Professional and Conviction Integrity Programs: Why We Need Them, Why They Will Work, and Models for Creating Them*, 31 CARDOZO L. REV. 2215, 2221 (2010) (internal quotation marks and internal citations omitted), which states that: The absolute immunity doctrine is not the only reason federal civil rights claims against prosecutors are, as a practical matter, rare and difficult to pursue. The qualified immunity good faith defense is a very substantial hurdle for a civil rights plaintiff as well. Though not a complete bar to liability, the Supreme Court has recognized that qualified immunity shields all but the plainly incompetent or those who knowingly violate the law. After *Ashcroft v. Iqbal*, it is certainly more difficult for a plaintiff, without any discovery, to file a pleading that will survive a motion to dismiss on qualified immunity grounds. Denials of the qualified immunity defense are also subject to interlocutory appeal, thereby making these lawsuits longer and more costly to litigate than ordinary cases. And even when the plaintiff prevails, there are serious problems collecting substantial damages awards.

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE..., 4 Va. J. Crim. L. 1

- 397 Interestingly, the doctrine of prosecutorial immunity and state bar discipline are connected. Amicus briefs filed by district attorneys and attorneys general groups advocating for the extension of prosecutorial immunity have claimed that the specter of such professional sanctions is a sufficient check. Scheck, *supra* note 396, at 2222 n.27.
- 398 *See id.* at 2222.
- 399 *See id.* at 2223-24.
- 400 *See id.*
- 401 NAS REPORT, *supra* note 4, at 23.
- 402 *Id.* at 214.
- 403 *Id.* at 26.
- 404 *Id.*
- 405 Hsu, *Convicted Defendants Left Uninformed of Forensic Flaws Found by Justice Department*, *supra* note 76. For a listing of affected cases, see *Convictions Linked to FBI Lab's Suspect Forensics*, WASH. POST, <http://www.washingtonpost.com/wp-srv/special/local/fbi-crime-lab-case-reviews/> (last visited Nov. 21, 2015).
- 406 DEPT. OF JUSTICE, AN ASSESSMENT OF THE 1996 DEPARTMENT OF JUSTICE TASK FORCE REVIEW OF THE FBI LABORATORY i (2014) [hereinafter TASK FORCE REVIEW].
- 407 *Id.*
- 408 *Id.*
- 409 *Id.* at ii.
- 410 *Id.* According to the report, an “independent scientist who later reviewed the case found the FBI Lab analysis to be scientifically unsupportable and the testimony overstated and incorrect.” *Id.*
- 411 *See* Hsu, *Convicted Defendants Left Uninformed of Forensic Flaws Found by Justice Department*, *supra* note 76 (“In one Texas case, Benjamin Herbert Boyle was executed in 1997, more than a year after the Justice Department began its review. Boyle would not have been eligible for the death penalty without the FBI's flawed work, according to a prosecutor's memo.”) Two others were executed prior to their cases being reviewed (though there was no finding of materiality); another capital defendant died in prison of natural causes before his case was reviewed. *See* TASK FORCE REVIEW, *supra* note 406, at ii.
- 412 *Id.* at iii. The report concluded that, of the 402 cases reviewed for the report, only in 13 were disclosures made to defendants or their last counsel of record. *Id.*
- 413 A 2013 investigation by the Associated Press revealed that at least twenty-four innocent men whose convictions and/or indictments were obtained through the use of bitemark **evidence** have been exonerated since 2000. *See* Myers, *Once Key in Some Cases, Bite Mark Evidence Now Derided as Unreliable*, *supra* note 81. Based on “decades of court records, archives, news reports” and interviews with “[t]wo dozen forensic scientists and other experts ... including some who had never before spoken to a reporter about their work,” the Associated Press' investigation was regarded as the “most comprehensive” audit of bitemark cases ever undertaken. *Id.*
- 414 *See, e.g.,* *Burke v. Town of Walpole*, 405 F.3d 66, 73 (1st Cir. 2005); *Stinson v. City of Milwaukee*, No. 09-C-1033, 2013 WL 5447916, at *1 (E.D. Wis. Sept. 30, 2013).
- 415 “The Biter,” i.e., the individual responsible for the bite mark at issue, was the highest level of certainty sanctioned by the ABFO until August of 2013, when the Reference Manual was updated. *See* DIPLOMATES REFERENCE MANUAL,

supra note 181, at 119. See also AM. BOARD OF FORENSIC ODONTOLOGY, GUIDELINES AND STANDARDS DRAFT 14 (2014) (“The ABFO does not support a conclusion of ‘The Biter’ in an open population case(s)”). Additionally, according to an article in the Wall Street Journal, although current president-elect of the ABFO, Dr. Peter Loomis, stated in July of this year that bite mark evidence could be used to “‘include or exclude’ a suspect,” he acknowledged that it “shouldn’t be used to identify a suspect.” Jack Nicas, *Flawed Evidence Under a Microscope: Disputed Forensic Techniques Draw Fresh Scrutiny; FBI Says It Is Reviewing Thousands of Convictions*, WALL ST. J. (July 18, 2013), <http://www.wsj.com/articles/SB10001424127887324263404578614161262653>.

416 In a recent *New York Times* article about Eddie Lee Howard’s case in Mississippi, the current president of the ABFO was quoted as saying that “actually naming an individual biter to a reasonable degree of certainty should be very limited.” Erik Eckholm, *Mississippi Death Row Case Faults Bite-Mark Forensics*, N.Y. TIMES (Sept. 15, 2014), <http://www.nytimes.com/2014/09/16/us/mississippi-death-row-appeal-highlights-shortcomings-of-bite-mark-identifications.html>. In Howard’s death penalty case, the ABFO member who testified--Dr. Michael West--stated “to a reasonable medical certainty” that Howard’s teeth inflicted the bite mark on the victim and further said, “Do I have any doubt [Howard’s] teeth made that bite on [the victim’s] breast? I don’t have any.” See Transcript of Record at 561, 584, State v. Howard, No. 92-400-CR1 (Lowndes Cnty. Cir. Ct. May 22, 2000).

417 In fact, in response to the *New York Times* article on Howard’s case, the ABFO posted the following on its website: *The New York Times* printed an article on 9/16/2014 faulting “bitemark forensics.” It highlights an appeal recently filed by the Mississippi Innocence Project with the Mississippi Supreme Court, a 22 year old case in which bite mark testimony was provided by Dr. Michael West. Like every news article, there are misstatements and some erroneous information is given. In particular, the author parrots the flawed Innocence Project publicity that 17 people previously convicted based on “expert bite matches” have been exonerated by DNA evidence. The IP often uses the number of 24 so at least the number is down a bit, but in actuality the number is 10, and of these, five of the opinions were not “match” as the article mentions but a lesser opinion. While any number of wrongful convictions is unacceptable and we are all cognizant of the fact that some terrible mistakes have been made in the past, we cannot ignore the fact that hundreds of positive outcomes have occurred throughout the country wherein bite mark evidence played a crucial role in the judicial process to assist the triers of fact. The ABFO continues to make changes to ensure accuracy of expert opinions. The ABFO has developed the Bitemark Analysis and Comparison Decision Tree, is continuing to develop a bitemark proficiency examination, has significantly raised the bitemark and other requirements for examination eligibility for new candidates, requires recertifying diplomates to take a recertification examination and has revised the standards, guidelines and terminology for bitemark analysis. Peter Loomis, President, Am. Board of Odontology, Third Quarter Message (Sept. 24, 2014).

418 See *supra* notes 274-75 and accompanying text.

419 See *id.*

420 A review of transcripts from state hair comparison cases during this period revealed a pattern of similar, invalid testimony by state hair experts, many of whom--if not most of whom--learned to provide such testimony at the FBI training course. See Post-Conviction Memorandum, *supra* note 385, at 2 n.2 (“Among other things, this affidavit discusses the FBI’s training; in it, Mr. Howard states: ‘I was taught at the FBI class that the best basis for testimony was our own experience through case-work.’ The affidavit relates to improper hair comparison testimony provided by one time chief of the Montana State Crime Laboratory Arnold Melnikoff in Jimmy Bromgard’s trial for raping a young girl. After nearly 15 years in prison, Mr. Bromgard was exonerated through post-conviction DNA testing.”).

421 The American Society of Crime Laboratory Directors/Laboratory Accreditation Board (ASCLD/LAB) recognized the need for possible state reviews, as well. ASCLD/LAB recommended “each laboratory, in consultation with the appropriate legal authorities, consider whether there may be past cases, specifically involving convictions, in which it would be appropriate to evaluate the potential impact of the reported conclusions and/or related testimony on the conviction.” ASCLD Press Release, *supra* note 382.

- 422 See generally CTR. ON THE ADMIN. OF CRIMINAL LAW, N.Y.U. SCHOOL OF LAW, ESTABLISHING CONVICTION INTEGRITY PROGRAMS IN PROSECUTOR'S OFFICES, available at http://www.law.nyu.edu/sites/default/files/upload_documents/Establishing_Conviction_Integrity_Programs_FinalReport_ecm_pro_073583.pdf.
- 423 See Bruce A. Green & Ellen Yaroshefsky, *Prosecutorial Discretion and Post-Conviction Evidence of Innocence*, 6 OHIO ST. J. CRIM. L. 467, 511 (2009).
- 424 See MODEL RULES OF PROF'L RESPONSIBILITY EC 7-13 (1980); see also AM. BAR ASS'N, AMERICAN BAR ASSOCIATION PROJECT ON STANDARDS FOR CRIMINAL JUSTICE, Standards 3-3.9(c), 3-3.11, available at http://www.americanbar.org/publications/criminal_justice_section_archive/crimjust_standards_pfunc_toc.html.
- 425 See Green & Yaroshefsky, *supra* note 423, at 482 nn.87 & 89, 483 n.96.
- 426 The cooperation of analysts may require that these individuals be accorded some limited immunity, either upon request or when set forth as a reason for not participating in a review.
- 427 See *supra* notes 139-42 and accompanying text; see also *infra* note 428.
- 428 Letter from John Crabb Jr., U.S. Dep't of Justice, to Robert P. McCulloch, St. Louis Cnty. Prosecutor's Office (Aug. 20, 2013) (on file with authors). There is ample precedent for this position. See, e.g., *Wilson v. Beard*, 426 F.3d 653, 661 (3d Cir. 2005) (holding that due diligence did not require prisoner to monitor local news twelve years after conviction when there was no reasonable basis to conclude that local news would provide information on prisoner's case); *United States v. Atchison*, No. 09-C-2105, 2012 WL 581163, at *1, 5 (N.D. Ill. Feb. 22, 2012) (holding that due diligence does not require prisoners "to hunt through haystacks trying to figure out whether one of them might contain a needle"); *Poole v. Woods*, No. 08-cv-12955, 2011 WL 4502372, at *1 (E.D. Mich. Aug. 9, 2011) (holding that due process claim based on discovery of faulty bite mark evidence was timely under applicable limitations period of AEDPA and that reasonable diligence did not require the "Petitioner to regularly scour the Detroit Free Press and Michigan Court Reporters more than a half-decade after his direct appeal was exhausted on the off chance that something unforeseeable yet useful to his case would be found"). In fact, most post-conviction petitioners have limited access--if any--to technical, scientific research. See generally *In re Personal Restraint of Trapp*, No. 65393-8-I, 2011 WL 5966266, at *1, 5 (Wash. Ct. App. Nov. 28, 2011) (holding that petition based on newly discovered CBLA evidence was not time barred because, while "a report generally calling CBLA evidence into question may have been published in 2004, the extent of the FBI's misleading testimony in [the petitioner's] case only became apparent after a detailed review of the trial record by specialists at the FBI laboratory sometime in 2009").
- 429 See, e.g., LA. CODE CRIM. PROC. ANN. art. 930.3(7) (2015). Some states allow newly discovered evidence arguments only in support of an actual innocence claim. See, e.g., MD. CODE ANN. § 8-301(a) (West 2015); CAL. PENAL CODE § 1473.6(a) (1) (West 2015) (newly discovered evidence must "point unerringly to his or her innocence").
- 430 See, e.g., *Giglio v. United States*, 405 U.S. 150, 154 (1972).
- 431 See, e.g., *Brecht v. Abrahamson*, 507 U.S. 619, 639 (1993) (Stevens, J., concurring) ("The Fourteenth Amendment prohibits the deprivation of liberty 'without due process of law'; that guarantee is the source of the federal right to challenge state criminal convictions that result from fundamentally unfair trial proceedings."); *Estelle v. McGuire*, 502 U.S. 62, 70 (1991) ("[T]he Due Process Clause guarantees fundamental elements of fairness in a criminal trial."); *Strickland v. Washington*, 466 U.S. 668, 696 (1984) ("[T]he ultimate focus of inquiry must be on the fundamental fairness of the proceeding whose result is being challenged. In every case the court should be concerned with whether, despite the strong presumption of reliability, the result of the particular proceeding is unreliable because of a breakdown in the adversarial process that our system counts on to produce just results."); *United States v. Cronin*, 466 U.S. 648, 656 (1984) ("The right to the effective assistance of counsel is thus the right of the accused to require the prosecution's case to survive the crucible of meaningful adversarial testing."); *Chambers v. Mississippi*, 410 U.S. 284, 294 (1973) ("The right of an accused in a criminal trial to due process is, in essence, the right to a fair opportunity to defend against the State's accusations."); *Spencer v. Texas*, 385 U.S. 554, 563-64 (1967) ("Cases in this Court have long proceeded on the premise that the Due Process Clause guarantees the fundamental elements of fairness

THE SHIFTED PARADIGM: FORENSIC SCIENCE'S OVERDUE..., 4 Va. J. Crim. L. 1

in a criminal trial.”); accord [Michigan v. Bryant](#), 562 U.S. 344, 370 n.13 (2011) (“The Due Process Clauses of the Fifth and Fourteenth Amendments may constitute a further bar to admission of, for example, unreliable **evidence**.”).

432 [Hyde v. State](#), 413 So. 2d 1042, 1044 (Miss. 1982).

433 *See id.* at 1043.

434 *See id.*

435 *See id.*

436 *See id.*

437 *Id.*

438 *See* Hsu, *Convicted Defendants Left Uninformed of Forensic Flaws Found by Justice Department*, *supra* note 76.

439 Letter from Amy Jabloner, U.S. Dep't of Justice, to Ben Saucier, Dist. Attorney, Jackson County (July 26, 2001) (on file with authors).

440 *Id.*

441 *See* Transcript of Record at 160, *State v. Hyde*, No. 53424 (Jackson Cnty. Cir. Ct. Apr. 24, 1982).

442 *Id.* at 170-71.

443 Fax from Ben Saucier, Dist. Attorney, Jackson County, to Ellis Gordon, U.S. Dept. of Justice (March 23, 2002).

4 VAJCL 1

End of Document

© 2017 Thomson Reuters. No claim to original U.S. Government Works.