

TRIAL JUDGES AND THE FORENSIC SCIENCE PROBLEM

STEPHANIE L. DAMON-MOORE*

In the last decade, many fields within forensic science have been discredited by scientists, judges, legal commentators, and even the FBI. Many different factors have been cited as the cause of forensic science's unreliability. Commentators have gestured toward forensic science's unique development as an investigative tool, cited the structural incentives created when laboratories are either literally or functionally an arm of the district attorney's office, accused prosecutors of being overzealous, and attributed the problem to criminal defense attorneys' lack of funding, organization, or access to forensic experts.

But none of these arguments explain why trial judges, who have an independent obligation to screen expert testimony presented in their courts, would routinely admit evidence devoid of scientific integrity. The project of this Note is to understand why judges, who effectively screen evidence proffered by criminal defendants and civil parties, fail to uphold their gatekeeping obligation when it comes to prosecutors' forensic evidence, and how judges can overcome the obstacles in the path to keeping bad forensic evidence out of court.

INTRODUCTION	1533
I. THE FORENSIC SCIENCE PROBLEM.....	1537
II. PROSECUTORS, DEFENDERS, AND JUDGES: WHO CAN FIX THE FORENSIC SCIENCE PROBLEM?.....	1541
A. <i>Prosecutors</i>	1541
B. <i>Criminal Defense Attorneys</i>	1546
C. <i>Judges</i>	1550
III. UNDERSTANDING THE JUDICIAL FAILURE TO GATEKEEP.....	1555
A. <i>Understanding How Judges Overlook Problems with Forensic Science</i>	1555
1. <i>Judges Lack Scientific Knowledge</i>	1555
2. <i>Judges Do Not Receive Helpful Information from Defense Attorneys</i>	1558
3. <i>Judges Rely on Heuristics to Admit Prosecution Evidence</i>	1559
B. <i>From Recognition to Action</i>	1563

* Copyright © 2017 by Stephanie L. Damon-Moore. J.D., 2017, New York University School of Law; B.A., 2011, Vassar College. Thank you to Judge Alex Kozinski for suggestions early in the process, Judge Harry T. Edwards for helpful comments and criticism, and Professor Erin Murphy for advice and support. I am also grateful for the early enthusiasm and ongoing diligence of the *N.Y.U. Law Review* editors, particularly Nick Krafft, Eliana Theodorou, Matthew Tieman, and Jiyae Hwang.

1. *Why Judges May Choose Not to Exclude Bad Evidence* 1563

2. *The Path to Solving the Forensic Science Problem* 1566

CONCLUSION 1570

INTRODUCTION

In the last decade, many fields within forensic science have been discredited by scientists,¹ judges,² legal commentators,³ and even the FBI.⁴ The foundation, methodology, execution, and conclusions of forensic scientists have been repeatedly called into question.⁵ Criticisms have been directed at all forensic disciplines, from bite-mark

¹ *E.g.*, PRESIDENT’S COUNCIL OF ADVISORS ON SCI. & TECH., FORENSIC SCIENCE IN CRIMINAL COURTS: ENSURING SCIENTIFIC VALIDITY OF FEATURE-COMPARISON METHODS (2016) [hereinafter PCAST REPORT], https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf (report authored by “an advisory group of the Nation’s leading scientists and engineers”).

² *E.g.*, Gimenez v. Ochoa, 821 F.3d 1136, 1144 & n.4 (9th Cir. 2016) (Kozinski, J.) (recognizing that we are in an “age where forensics that were once considered unassailable are subject to serious doubt” and collecting sources); Harry T. Edwards & Jennifer L. Mnookin, Opinion, *A Wake-Up Call on the Junk Science Infesting Our Courtrooms*, WASH. POST (Sept. 20, 2016), https://www.washingtonpost.com/opinions/a-wake-up-call-on-the-junk-science-infesting-our-courtrooms/2016/09/19/85b6eb22-7e90-11e6-8d13-d7c704ef9fd9_story.html (opinion piece by a D.C. Circuit Judge and the Dean of UCLA School of Law); Nancy Gertner, Opinion, *Judges Need to Set a Higher Standard for Forensic Evidence*, N.Y. TIMES: ROOM FOR DEBATE (Feb. 4, 2016, 6:36 PM), <http://www.nytimes.com/roomfordebate/2015/03/30/robert-durst-handwriting-and-judging-forensic-science/judges-need-to-set-a-higher-standard-for-forensic-evidence> (editorial authored by a former federal district court judge).

³ *See, e.g.*, Keith A. Findley, *Reforming the ‘Science’ in Forensic Science*, 88 WISC. LAW. 32 (2015); Brandon L. Garrett & Peter J. Neufeld, *Invalid Forensic Science Testimony and Wrongful Convictions*, 95 VA. L. REV. 1 (2009); Jennifer L. Mnookin, *The Validity of Latent Fingerprint Identification: Confessions of a Fingerprinting Moderate*, 7 LAW, PROBABILITY & RISK 127 (2008).

⁴ *See, e.g.*, Press Release, FBI, FBI Testimony on Microscopic Hair Analysis Contained Errors in at Least 90 Percent of Cases in Ongoing Review (Apr. 20, 2015) [hereinafter Microscopic Hair Analysis Press Release], <https://www.fbi.gov/news/pressrel/press-releases/fbi-testimony-on-microscopic-hair-analysis-contained-errors-in-at-least-90-percent-of-cases-in-ongoing-review> (quoting Peter Neufeld, Co-Director of the Innocence Project, concluding that “FBI microscopic hair analysts committed widespread, systematic error, grossly exaggerating the significance of their data under oath with the consequence of unfairly bolstering the prosecutions’ case”).

⁵ *See* NAT’L RESEARCH COUNCIL, STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD (2009) [hereinafter NRC REPORT], <https://www.ncjrs.gov/pdffiles1/nij/grants/228091.pdf>; PCAST REPORT, *supra* note 1; Craig M. Cooley, *Nurturing Forensic Science: How Appropriate Funding and Government Oversight Can Further Strengthen the Forensic Science Community*, 17 TEX. WESLEYAN L. REV. 441 (2011); Garrett & Neufeld, *supra* note 3.

analysis⁶ to DNA analysis.⁷ Hundreds if not thousands of convictions that rested in whole or in part on this “junk science” have been overturned,⁸ while innumerable additional convictions have been cast into doubt. Yet this evidence has been and continues to be routinely admitted in criminal court.⁹

Commentators have pointed to a number of possible explanations for the problems with forensic science evidence: forensic science’s unique development as an investigative tool;¹⁰ the structural incentives created when laboratories are either literally or functionally an arm of the district attorney’s office;¹¹ and inadequate education programs for forensics.¹² They have blamed the proliferation of

⁶ See Mary A. Bush et al., *Statistical Evidence for the Similarity of the Human Dentition*, 56 J. FORENSIC SCI. 118 (2011) (debunking a foundational assumption of bite-mark analysis that human mouths are unique); see also PCAST REPORT, *supra* note 1, at 83–87 (surveying the limited research on bite-mark analysis and concluding that “the observed false positive rates were so high that the method is clearly scientifically unreliable at present”).

⁷ See Matthew Shaer, *The False Promise of DNA Testing*, ATLANTIC, June 2016, <https://www.theatlantic.com/magazine/archive/2016/06/a-reasonable-doubt/480747/> (describing how DNA testing has become less reliable as scientists rely on increasingly smaller samples and larger numbers of contributors); see also ERIN E. MURPHY, *INSIDE THE CELL: THE DARK SIDE OF FORENSIC DNA* 29–150 (2015) (describing problems with both DNA analysis and testimony by DNA analysts).

⁸ *Misapplication of Forensic Science*, INNOCENCE PROJECT, <https://www.innocenceproject.org/causes/misapplication-forensic-science/> (last visited July 7, 2017).

⁹ See, e.g., NRC REPORT, *supra* note 5, at 106 (“Review of reported judicial opinions reveals that, at least in criminal cases, forensic science evidence is not routinely scrutinized pursuant to the [applicable] standard of reliability”); Erin Murphy, *Neuroscience and the Civil/Criminal Daubert Divide*, 85 FORDHAM L. REV. 619, 624 (2016) (“[N]early all of the common forensic techniques . . . routinely admitted by courts[] have been repeatedly denounced as lacking in any scientific basis.”).

¹⁰ See, e.g., NRC REPORT, *supra* note 5, at 42 (“Most [forensic] techniques were developed in crime laboratories to aid in the investigation of evidence from a particular crime scene, and researching their limitations and foundations was never a top priority.”).

¹¹ See, e.g., David E. Bernstein, *Expert Witnesses, Adversarial Bias, and the (Partial) Failure of the Daubert Revolution*, 93 IOWA L. REV. 451, 460 (2008) (noting that forensic labs’ frequent dependence on police departments for their budgets may lead to a “desire to please the police”); Glen Whitman & Roger Koppl, *Rational Bias in Forensic Science*, 9 LAW, PROBABILITY & RISK 69, 71 (2010) (finding that “tension exists within an institutional structure that frequently puts crime labs under the administration of law enforcement agencies”).

¹² See, e.g., NRC REPORT, *supra* note 5, at 224–25 (“Commentators have noted repeatedly the deficiencies of forensic science education programs.”); KELLY M. PYREK, *FORENSIC SCIENCE UNDER SIEGE: THE CHALLENGES OF FORENSIC LABORATORIES AND THE MEDICO-LEGAL DEATH INVESTIGATION SYSTEM* 93–96 (2007) (“[A] chronic lack of funding keeps many forensic practitioners from the pursuit of much-needed training.”); Cooley, *supra* note 5, at 450–60 (noting that many forensic analysts are “inadequately trained” in the relevant sciences); Joseph L. Peterson & Anna S. Leggett, *The Evolution of Forensic Science: Progress amid the Pitfalls*, 36 STETSON L. REV. 621, 626 (2007) (noting that studies have demonstrated a “severe need for higher-education programs to prepare future forensic scientists for positions in government laboratories”).

invalid forensic science testimony in the courtroom on overzealous prosecutors¹³ and criminal defense attorneys' incompetence, lack of funding, and lack of access to forensic experts.¹⁴

But none of these explanations account for why trial judges, who have an independent obligation to screen expert testimony presented in their courts, routinely admit evidence and permit testimony devoid of scientific integrity. And given the rigor with which trial judges screen experts proffered by criminal defendants and civil litigants, no serious argument can be made that judges lack the scientific savvy to execute their duty responsibly.¹⁵ Nor, for that matter, are the shortcomings of forensic science especially complex.¹⁶ Yet forensic evidence has been treated as reliable, even when the experts themselves acknowledge glaring gaps in their scientific methods and conclusions.¹⁷

Although scholars have noted that trial judges do not rigorously uphold their expert admissibility screening obligation when it comes to prosecutors' forensic experts, few have tried to explain why.¹⁸ One

¹³ See Stanley Z. Fisher, *In Search of the Virtuous Prosecutor: A Conceptual Framework*, 15 AM. J. CRIM. L. 197, 208–11 (1988) (examining structural incentives that push prosecutors to emphasize convictions).

¹⁴ See, e.g., Peter J. Neufeld, *The (Near) Irrelevance of Daubert to Criminal Justice and Some Suggestions for Reform*, 95 AM. J. PUB. HEALTH S107, S110 (2005); see also *infra* Section II.B (discussing challenges defense attorneys face in challenging bad forensic evidence proffered by prosecutors).

¹⁵ See *infra* notes 145–46; see also Neufeld, *supra* note 14, at S109 (reviewing criminal and civil decisions and finding that judges excluded evidence proffered by civil plaintiffs and criminal defendants much more often than evidence proffered by prosecutors). *But see* Jennifer L. Groscup & Steven D. Penrod, *Battle of the Standards for Experts in Criminal Cases: Police vs. Psychologists*, 33 SETON HALL L. REV. 1141, 1144–46 (2003) (surveying the literature on judicial understanding of *Daubert* factors and finding that judges may not understand them all).

¹⁶ See *infra* Part I (describing some of the most basic and most common problems with forensic evidence).

¹⁷ See, e.g., *infra* notes 31–34 and accompanying text (discussing testimony by bite-mark experts).

¹⁸ See, e.g., Findley, *supra* note 3, at 39 (noting that “federal courts applying *Daubert* almost never exclude prosecution-proffered forensic science evidence,” but do exclude evidence proffered by criminal defendants and civil plaintiffs, and that the reason for this disparate treatment “is not fully understood”); Jane Campbell Moriarty, “*Misconvictions, Science, and the Ministers of Justice*,” 86 NEB. L. REV. 1, 37 (2007) (noting that “trial courts have steadfastly refused to take the *Daubert* trilogy language seriously as applied to [certain forensic] evidence,” without addressing why that is the case); Jane Campbell Moriarty & Michael J. Saks, *Forensic Science: Grand Goals, Tragic Flaws, and Judicial Gatekeeping*, JUDGES’ J., Fall 2005, at 16, 28–29 (noting that judges fail to uphold gatekeeping responsibilities in both *Frye* and *Daubert* jurisdictions, without taking up the source of this failure); Michael J. Saks & Ashley M. Votruba, “. . . *And the Courts Have Been Utterly Ineffective*,” JUDGES’ J., Summer 2015, at 28, 29–30 (same). *But see* Julie A. Seaman, *A Tale of Two Dauberts*, 47 GA. L. REV. 889, 900, 902 (2013) (suggesting that variation in the quality of experts and, reluctance in criminal cases, to “exclude prosecution

notable exception is retired Judge Donald Shelton of Michigan's 22nd Circuit Court in Ann Arbor, Michigan, currently the director of the Criminal Justice Studies Program at the University of Michigan-Dearborn.¹⁹ Judge Shelton has argued that judicial permissiveness with regard to forensic evidence is most reasonably attributed to judicial bias—or “attitudinal blinders”—and has suggested some of the same causes that I will discuss, including the fact that many judges are former prosecutors, as well as the influence of “tough on crime” elections.²⁰ Although I agree with Judge Shelton that judicial bias likely influences expert admissibility decisions,²¹ my analysis recognizes a number of constraints on judges' abilities to recognize and address problems with forensic science, including judges' lack of scientific training,²² the frequent absence of defense objections to prosecution proffers,²³ and concerns—rooted in both law and policy—about withholding relevant evidence from juries.²⁴

For the purposes of this Note, I define the “forensic science problem” as the admission of unreliable forensic testimony in criminal trials. My framing reflects the view that the worst outcome of “junk science” is wrongful convictions,²⁵ and accordingly that preventing

evidence that carries a long historical pedigree” could explain admission of handwriting analysis evidence in criminal cases at higher rates than civil cases).

¹⁹ *Criminology & Criminal Justice Studies*, UNIV. MICH.-DEARBORN, <https://umdearborn.edu/casl/undergraduate-programs/areas-study/criminology-criminal-justice-studies> (last visited July 9, 2017).

²⁰ Donald E. Shelton, *Forensic Science Evidence and Miscarriages of Justice*, in *ADVANCES IN FORENSIC HUMAN IDENTIFICATION* 409, 417–21 (Xanthé Mallett et al. eds., 2014) [hereinafter Shelton, *Miscarriages of Justice*]; Donald E. Shelton, *Forensic Science Evidence and Judicial Bias in Criminal Cases*, *JUDGES' J.*, Summer 2010, at 18, 22–24 [hereinafter Shelton, *Judicial Bias*]; see also Adam B. Shniderman, *Prosecutors Respond to Calls for Forensic Science Reform: More Sharks in Dirty Water*, 126 *YALE L.J.F.* 348, 353–56 (2017) (attributing “the judiciary’s failure to keep bad science out of courtrooms” to “a systemic pro-prosecution bias on the bench”).

²¹ See *infra* notes 157–71 (discussing the influence of biases and heuristics on admissibility decisions).

²² See *infra* Section III.A.1 (discussing the problem that judges commonly lack scientific training and may struggle to rigorously assess reliability of expert evidence as a result).

²³ See *infra* notes 149–56 and accompanying text (discussing the relative absence of effective defense challenges to prosecution experts).

²⁴ See *infra* notes 173–85 and accompanying text (describing how the Federal Rules of Evidence as well as broader policy concerns may lead judges to err on the side of admitting untested evidence).

²⁵ As Blackstone said, and many a legal commentator has echoed, it is “better that ten guilty persons escape, than that one innocent suffer.” 4 WILLIAM BLACKSTONE, *COMMENTARIES* *352. Whether we can reasonably infer anything from the widespread adoration of this sentiment in the legal community is an entirely different question. See generally Alexander Volokh, n *Guilty Men*, 146 *U. PA. L. REV.* 173 (1997) (providing a cheeky analysis of the various versions of this sentiment).

wrongful convictions is the highest priority for reform. Others may define the worst outcome as the failure to convict the correct perpetrator, allowing a guilty person to remain “on the streets.” This definition, in turn, may prompt a focus on solutions such as increased federal funding for forensic science, or improvements to forensic science education programs. My definition, of course, points to a different kind of solution: keeping bad evidence out.

This Note proceeds as follows: Part I summarizes a few of the most common—and problematic—shortcomings of forensic science evidence. Part II examines the relative abilities and incentives of prosecutors, defense attorneys, and judges to solve the forensic science problem, and concludes that only judges are well positioned to do so. Lastly, Part III addresses the big question: How can judges uphold their gatekeeping responsibilities and keep junk science out of the courtroom?

I

THE FORENSIC SCIENCE PROBLEM

In recent years, nearly all forensic disciplines have been criticized as insufficiently tested or even invalid.²⁶ Though a comprehensive analysis of the problems with various forensic disciplines is beyond the scope of this Note,²⁷ it is worth fleshing out some of these issues, to both bring the forensic science problem to life and to dispel the notion that there is something special about forensic evidence that makes it especially difficult to screen out. A brief foray into the character of the problems with forensic evidence makes it clear that the admission of flawed forensic science cannot be so easily explained.

The paradigmatic forensic science evidence involves a comparison of two samples—one taken from a known suspect, the other found at the scene of the crime—and the conclusion that there is a “match.”²⁸ Whether through the statements of the witness or the

²⁶ The only forensic science discipline that has generally escaped criticism is single-source and simple-mixture DNA testing. Compare PCAST REPORT, *supra* note 1 (finding serious flaws across nearly all disciplines evaluated, including DNA analysis of complex-mixture samples, bite-mark analysis, latent fingerprint analysis, firearms analysis, footwear analysis, and hair analysis), *with id.* at 69–73 (describing single-source and simple-mixture DNA analysis as “objective methods whose foundational validity has been properly established,” but cautioning that these methodologies are “not infallible in practice”).

²⁷ For detailed analyses of the shortcomings of forensic evidence, see generally PCAST REPORT, *supra* note 1, and NRC REPORT, *supra* note 5. See also Garrett & Neufeld, *supra* note 3 (cataloguing problems with forensic testimony specifically).

²⁸ PCAST actually advises against permitting experts to testify to a “match,” as the term is “likely to imply an inappropriately high probative value.” PCAST REPORT, *supra* note 1, at 45–46. They recommend the use of the phrase “*proposed* identification” instead. *Id.* at 46.

surmise of the jury, this match is interpreted to demonstrate with virtual certainty that the defendant was at the scene of the crime. The validity of this conclusion rests on two interlocking assumptions: first, that a match is defined by some kind of industry standard, as opposed to the personal view of the testifying expert, and second, that this match is inculpatory because the particular combination of features that matched the exemplar is either literally unique or at least highly unusual.

Research, however, has shown that these necessary assumptions are not present in many forensic fields. Several forensic disciplines have been criticized for their lack of an industry-wide match standard. For example, bite-mark analysis, also called forensic odontology, has no industry-wide standard defining how many points of similarity must exist to call a comparison a “match.” Instead, the reference manual for the American Board of Forensic Odontology’s “Human Bitemark Analysis Guidelines” is comprised of a cursory, bullet-point-style list of features to be documented, along with suggestions as to how the bite mark could be compared to exemplars.²⁹ Although the guidelines rather unhelpfully list some vocabulary for articulating whether a bite mark does or does not match an exemplar, notably absent are any *standards* for determining whether said match exists.³⁰ In other words, bite-mark “experts” have some parameters for the types of criteria to consider and the way to articulate their findings, but the critical middle step—drawing the correct conclusions from the evidence—is left to individual discretion.

Remarkably, forensic odontologists have not been shy about the fact that their “scientific” field is entirely devoid of standards for matching samples.³¹ For example, in a leading case from the Missouri Court of Appeals, *State v. Sager*, the bite-mark expert acknowledged that there was “no standard procedure for arriving at conclusions” in the field of forensic odontology, and that his methodology “would not

²⁹ See AM. BD. OF FORENSIC ODONTOLOGY, DIPLOMATES REFERENCE MANUAL 98–99 (Mar. 2017 ed.), <http://abfo.org/wp-content/uploads/2012/08/ABFO-Reference-Manual-April-2017-v7.pdf>.

³⁰ See *id.* at 102.

³¹ See 4 DAVID L. FAIGMAN ET AL., MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY § 35:4 (2016–2017 ed. 2016) (noting the irony that although “forensic odontologists, perhaps reflecting a grounding in scientific skepticism that is often absent from the more traditional forensic identification sciences, were themselves somewhat doubtful about whether the state of their knowledge permitted them to successfully identify a perpetrator ‘to the exclusion of all others,’” admission by courts apparently convinced them “that, despite their doubts, they were indeed able to perform bitemark identifications after all” (internal citation omitted)).

necessarily be used by all experts.”³² Along the same lines, Mississippi’s Supreme Court has repeatedly recognized that there “is little consensus in the scientific community on the number of points which must match before any positive identification can be announced” and that “methods of comparison employed in a particular case may differ.”³³

Yet, in these and other cases, the evidence gets admitted. In *Sager*, for example, the court not only allowed the expert to testify that the bite mark “matched,” but also allowed him to testify that “no two people could have identical mouths,” that “the bite mark reflected in the photograph of the breast of the victim was *beyond a reasonable doubt* placed upon the victim’s breast by appellant,” and that “the conclusions reached [by all experts] would be the same.”³⁴ This confidence is misplaced. In the limited scientific studies of bite-mark analysis, forensic odontologists disagreed not only about whether there was a match to the exemplar, but even about whether the mark was made by a human.³⁵

Bite-mark analysis may be the poster child for a forensic field that commentators love to hate,³⁶ but another forensic discipline that lacks standards for a “match” is a former darling of the forensic evidence world: fingerprint analysis.³⁷ The FBI once proudly wrote that “[o]f all the methods of identification, fingerprinting alone has proved to be both infallible and feasible.”³⁸ But while courts have been

³² 600 S.W.2d 541, 563 (Mo. Ct. App. 1980) (admitting bite-mark analysis testimony despite lack of standardized methodology).

³³ *Stubbs v. State*, 845 So. 2d 656, 669 (Miss. 2003) (quoting *Howard v. State*, 701 So. 2d 274, 288 (Miss. 1997)).

³⁴ *Sager*, 600 S.W.2d at 563 (emphasis added). Though the claim that human mouths are unique goes to the heart of forensic odontology, that claim was debunked in a 2010 study. Bush et al., *supra* note 6.

³⁵ See PCAST REPORT, *supra* note 1, at 84–85.

³⁶ In addition to the pernicious problems in forensic odontology, commentators have also had a field day with bite-related puns. *E.g.*, Erica Beecher-Monas, *Reality Bites: The Illusion of Science in Bite-Mark Evidence*, 30 CARDOZO L. REV. 1369 (2009); Adam Deitch, Comment, *An Inconvenient Tooth: Forensic Odontology Is an Inadmissible Junk Science When It Is Used to “Match” Teeth to Bitemarks in Skin*, 2009 WISC. L. REV. 1205.

³⁷ See Lyn Haber & Ralph Norman Haber, *Scientific Validation of Fingerprint Evidence Under Daubert*, 7 LAW, PROBABILITY & RISK 87, 105 (2008) (finding no support for the scientific validity of the most common method of fingerprint analysis); Mnookin, *supra* note 3, at 131 (noting the “near-universal” judicial acceptance of fingerprint analysis, and her own conclusion that most fingerprint evidence should be excluded under *Daubert*). Of course, today the “gold standard” forensic discipline is DNA evidence, which has recently garnered some criticism of its own. See *supra* note 7; see also PCAST REPORT, *supra* note 1, at 78–86 (describing problems with DNA analysis of complex-mixture samples).

³⁸ FBI, THE SCIENCE OF FINGERPRINTS: CLASSIFICATIONS AND USES iv (1984).

admitting fingerprint evidence for over a hundred years,³⁹ the discipline lacked an industry-wide match standard until the National Research Council criticized this shortcoming in a 2009 report.⁴⁰ Since then, the Scientific Working Group on Friction Ridge Analysis, Study and Technology has issued standards that, according to the President's Council of Advisors on Science and Technology (PCAST), "begin[] to move latent print analysis in the direction of an objective framework."⁴¹ Nevertheless, PCAST concluded that fingerprint analysis has "a considerable way to go" before achieving objectivity.⁴²

If and when a match *can* be reliably established in any forensic discipline, the significance of the match hinges on the degree to which a match is *unusual*. Consider, for example, blood types: A serologist might find that the defendant's blood type matches a sample found at the crime scene. How inculpatory that is depends precisely on how unusual the blood type is. For instance, eleven percent of the white population have Type B blood.⁴³ So, if you have a white Type B defendant, and you have a Type B sample taken from the scene, you can safely say that the sample excludes eighty-nine percent of the white population and does not exclude the defendant.

On the other hand, if you don't know anything about how common each blood type is, you could testify that you found a match, but not about how unusual that match is. Take, for example, microscopic hair analysis. We don't know how common it is for people within a racial group to have similar hair characteristics, which means that testimony about hair characteristics ought to be limited to a statement that the sample found at the scene "could have" come from the defendant.⁴⁴ Unfortunately, microscopic hair experts have struggled to restrain themselves: For decades, experts testified that only about "5 percent of the population" shared certain characteristics, or that finding a match put the odds of a false positive at "one chance in 10,000."⁴⁵ In 2015, an FBI review concluded that, of the twenty-eight

³⁹ See Mnookin, *supra* note 3, at 128 (describing judicial opinions that laud fingerprint evidence as having "survived an entire century of testing within the crucible of the courtroom").

⁴⁰ NRC REPORT, *supra* note 5, at 140–41 (discussing the degree to which latent fingerprint analysis relies on the subjective interpretation of individual examiners).

⁴¹ PCAST REPORT, *supra* note 1, at 91.

⁴² *Id.* at 88. The PCAST Report also concludes that "estimated false positive rates are *much higher* than the general public . . . would likely believe based on longstanding claims about the accuracy of fingerprint analysis." *Id.* at 95.

⁴³ Garrett & Neufeld, *supra* note 3, at 36 (citing Dale D. Dykes, *The Use of Frequency Tables in Parentage Testing*, in PROBABILITY OF INCLUSION IN PATERNITY TESTING: A TECHNICAL WORKSHOP 15, 20, 29 (Herbert Silver ed., 1982)).

⁴⁴ *Id.* at 49.

⁴⁵ *Id.* at 53–54.

FBI agents who conducted microscopic hair analyses, *twenty-six* made erroneous statements in written reports or oral testimony.⁴⁶

This brief description of a few common problems with forensic evidence demonstrates that the shortcomings of forensic science are not so complex that prosecutors cannot understand them, defense attorneys cannot challenge them, and judges cannot bar evidence because of them. Yet over and over again forensic expert witnesses are permitted to give invalid testimony. The following Part looks at the three primary actors within the criminal justice system—prosecutors, criminal defense attorneys, and trial judges—to determine who is best suited to address the forensic science problem.

II

PROSECUTORS, DEFENDERS, AND JUDGES: WHO CAN FIX THE FORENSIC SCIENCE PROBLEM?

A. Prosecutors

Prosecutors have a great degree of power to prevent erroneous admissions. As the actors who proffer the evidence in the first place, they could all but solve the forensic science problem tomorrow simply by declining to present evidence that lacks a valid foundation and a reliable methodology.⁴⁷ But, despite the expectation that prosecutors will act as “minister[s] of justice,”⁴⁸ the adversarial structure of the criminal justice system incentivizes zealous prosecutions.⁴⁹ Scholars have repeatedly noted the tension between the stated neutrality of the prosecutor and the reality of prosecution in the United States.⁵⁰ While

⁴⁶ Microscopic Hair Analysis Press Release, *supra* note 4.

⁴⁷ Such a decision would not only solve the problem that is the focus of this Note—wrongful convictions based on junk science—but, by eliminating demand for “junk science” and creating demand for reliable evidence, would also provide a massive incentive for the forensic science community to adopt scientific approaches to forensic testing, and abandon disciplines that could not meet high standards.

⁴⁸ MODEL RULES OF PROF'L CONDUCT r. 3.8 cmt. 1 (AM. BAR ASS'N 2014) (declaring that “[a] prosecutor has the responsibility of a minister of justice and not simply that of an advocate,” which requires prosecutors to “see that the defendant is accorded procedural justice,” ensure “that guilt is decided upon the basis of sufficient evidence,” and take “special precautions . . . to prevent and rectify the conviction of innocent persons”).

⁴⁹ See Fisher, *supra* note 13, at 208–11 (describing ways in which the adversarial system leads prosecutors to “invest [their] energies single-mindedly in maximizing convictions and punishments”); see also Stephanos Bibas, *Prosecutorial Regulation Versus Prosecutorial Accountability*, 157 U. PA. L. REV. 959, 985 & n.98 (2009) (describing how media coverage favors prosecutors with high conviction rates).

⁵⁰ See, e.g., Kenneth Bresler, Essay, “*I Never Lost a Trial*”: When Prosecutors Keep Score of Criminal Convictions, 9 GEO. J. LEGAL ETHICS 537, 541–44 (1996) (discussing how the practice of some prosecutors to tally their convictions violates “[t]he most notable and noble principle of prosecution . . . to seek justice, not convictions”); Fred C. Zacharias, *Structuring the Ethics of Prosecutorial Trial Practice: Can Prosecutors Do Justice?*, 44

a prosecutor's role is ostensibly to reach the truth, some prosecutors have evinced marked resistance to the truth when the truth happens to be that the defendant is innocent.⁵¹

States' experts are not immune from the conviction mindset, either, perhaps because of the ways in which forensic laboratories are aligned with law enforcement. For example, many forensic science disciplines were initially developed by law enforcement for investigative purposes.⁵² And in most jurisdictions, forensic laboratories are still literally a part of the prosecutors' office.⁵³ Even where there is no formal connection, forensic labs get the vast majority of their business from prosecutors and law enforcement, potentially creating a sense

VAND. L. REV. 45, 107–08 (1991) (arguing that the “do justice’ mandate” is hamstrung by the adversarial nature of criminal prosecutions where “[w]inning is at a premium” and “competitive juices flow”).

⁵¹ See, e.g., DEBORAH TUERKHEIMER, *FLAWED CONVICTIONS: ‘SHAKEN BABY SYNDROME’ AND THE INERTIA OF JUSTICE* 148 (2014) (describing a shaken baby syndrome (SBS) case where the prosecutor dismissed the charges after the state's own experts disavowed a shaken baby syndrome diagnosis, but stated the office “wholeheartedly” believed the defendant was guilty nonetheless); Shaila Dewan, *Despite DNA Test, a Case Is Retried*, N.Y. TIMES (Sept. 6, 2007), <http://www.nytimes.com/2007/09/06/us/06dna.html> (describing District Attorney Forrest Allgood's decision to retry a rape and murder case after the convicted man, Kennedy Brewer, was excluded by DNA); see also Radley Balko, *Opinion, Election Results: One of America's Worst Prosecutors Lost Last Night, but One of Its Worst Attorney Generals Won*, WASH. POST (Nov. 4, 2015), <https://www.washingtonpost.com/news/the-watch/wp/2015/11/04/election-results-one-of-americas-worst-prosecutors-lost-last-night-but-one-of-its-worst-attorneys-general-won> (noting that Kennedy Brewer spent an additional seven years in prison after DNA exonerated him because of District Attorney Forrest Allgood's decision to re-prosecute him for the crime).

⁵² SANDRA GUERRA THOMPSON, *COPS IN LAB COATS: CURBING WRONGFUL CONVICTIONS THROUGH INDEPENDENT FORENSIC LABORATORIES* 109 (2015) (describing the development of forensic science as a tool for investigating and prosecuting crimes). As an investigatory tool, untested or unreliable methods may still be helpful. See PCAST REPORT, *supra* note 1, at 4 (noting that forensic science used for investigative purposes, as opposed to prosecution purposes, may appropriately “come from both well-established science and exploratory approaches”). But untested methods that are worth a shot when investigating a case can return wrongful convictions if introduced in court.

⁵³ See NAT'L CONFERENCE OF STATE LEGISLATURES, ADMINISTRATION OF FORENSIC SCIENCE DEPARTMENTS 1 (2013), <http://www.ncsl.org/Documents/cj/AdministrationOfForensicServices.pdf> (noting that only six states—Alabama, Arkansas, Connecticut, Delaware, Rhode Island, and Virginia—and the District of Columbia have forensic laboratories “that operate as . . . independent agenc[ies] or as . . . department[s] . . . that do [] not share oversight with prosecutorial or law enforcement services”). For a comprehensive analysis of the ways in which law enforcement interests have obstructed efforts to fund and perform independent scientific testing of forensic science, see Paul C. Giannelli, *Daubert and Forensic Science: The Pitfalls of Law Enforcement Control of Scientific Research*, 2011 U. ILL. L. REV. 53.

that success on the job means finding a match and getting a conviction.⁵⁴

One example of the dangers of such a “team” mentality among prosecutors and experts comes from the area of Shaken Baby Syndrome (SBS). Despite being based on what can charitably be described as a thin medical foundation,⁵⁵ growing national awareness of child abuse, among other factors, led to an explosion in prosecutions for SBS in the United States in the nineties and the first decade of the twenty-first century.⁵⁶ Many of these prosecutions came out of child abuse units in hospitals and collaboration between prosecutors and child abuse physicians in cases where a caregiver was alleged to have shaken an infant, causing brain damage or death.⁵⁷ In her seminal book on SBS prosecutions, Professor Deborah Tuerkheimer describes the role physicians played in SBS cases: “Doctors came to court and explained that, notwithstanding the absence of any other signs of abuse, shaking could be proved by three neurological symptoms: bleeding beneath the outer layer of membranes surrounding the brain, bleeding in the retina, and brain swelling.”⁵⁸ These doctors would go on to testify that the shaking must have been the cause, that it must have been unreasonably violent shaking, and that the perpetrator must have been the last person with the infant.⁵⁹ The medical testimony thus wasn’t merely the most damning part of the case; it *was* the case.

Against that backdrop, the medical community writ large began to call these conclusions into question. In 2011, for example, an advisory board member of the National Center on Shaken Baby Syndrome described the three symptoms as a “myth,” dismissing the notion that any “trained pediatrician” could equate them with abuse.⁶⁰ But child abuse units full of physicians trained to make just

⁵⁴ See Brandon L. Garrett, *The Genetic Panopticon*, BOS. REV. (Mar. 17, 2016), <http://bostonreview.net/us/brandon-garrett-dna-crime-lab-forensics> (describing police and prosecutors as the “chief consumers of forensic evidence”).

⁵⁵ As Professor Deborah Tuerkheimer points out, the doctor whose paper served as the foundation for SBS, A. Norman Guthkelch, has been a vocal critic of SBS prosecutions, writing that there was “not a vestige of proof when the name [SBS] developed that shaking alone causes the triad [of symptoms said to characterize SBS].” TUERKHEIMER, *supra* note 51, at 87.

⁵⁶ See *id.* at 2 (noting that only fifteen SBS cases reached the appeals courts before 1990, more than two hundred were appealed between 1990 and 2000, and more than eight hundred were appealed between 2000 and 2010).

⁵⁷ *Id.* at 35–37 (discussing the development of child abuse units and the impact these units have on SBS prosecutions).

⁵⁸ *Id.* at xi.

⁵⁹ *Id.*

⁶⁰ *Id.* at 17.

such a conclusion were already in place. A study by Northwestern University that examined 3000 SBS cases from a twenty-five-year period exposed what is arguably the result of collaboration between prosecutors and physicians: SBS prosecution “hot spots.”⁶¹ For example, Sarpy County, Nebraska, and nearby Douglas County, Nebraska, are number one and number four respectively in SBS prosecutions per capita.⁶² These blockbuster numbers of prosecutions may reflect the local “team” strategy, in which “law enforcement, child advocacy centers, prosecutors, the Department of Health and Human Services, and medical professionals” work together to investigate possible child abuse.⁶³ As this example demonstrates, adversarial incentives influence not only prosecutors themselves, but may influence experts with whom they work closely as well.

Not surprisingly given the adversarial structure of the criminal justice system, prosecutors have shown little interest in policing their own evidence. The National District Attorneys Association (NDAA), for instance, responded to the publication of the 2016 PCAST Report⁶⁴ by criticizing the committee’s methods and composition and rejecting all of its findings. “Experience shows these disciplines offer reliable and powerful evidence in a court of law,”⁶⁵ the NDAA argued. The NDAA continued: “It is therefore entirely inappropriate for the report to suggest otherwise to this country’s courts.”⁶⁶ Attorney General Loretta Lynch was no more interested in adopting PCAST’s findings, despite being appointed by the same liberal administration that commissioned the report. Her response was that the Department of Justice “remain[s] confident that, when used properly, forensic science evidence helps juries identify the guilty and clear the innocent,” and “[w]hile we appreciate [PCAST’s] contribution to the field of scientific inquiry, the department will not be adopting the rec-

⁶¹ See Lauryn Schroeder, *Pinpointing Shaken-Baby Syndrome Cases: A New Medill Justice Project Study Identifies Where Higher Rates of Shaken-Baby Syndrome Cases Are Occurring in the United States*, MEDILL JUST. PROJECT (Dec. 10, 2013), <http://www.medilljusticeproject.org/2013/12/10/hot-spots/> (suggesting that these prosecution “hot spots” are, in part, the result of a collaboration among medical professionals and prosecutors).

⁶² *Id.*

⁶³ *Id.*

⁶⁴ PCAST REPORT, *supra* note 1.

⁶⁵ Letter from Michael A. Ramos, President, Nat’l Dist. Attorneys Ass’n, to President Barack Obama 8 (Nov. 16, 2016), <http://www.ciclt.net/ul/ndaajustice/PCAST/NDAA%20PCAST%20Response%20FINAL.pdf>. These are choice words from the NDAA given the PCAST Report’s emphatic argument that experience cannot be used as a basis to determine the validity of forensic disciplines. PCAST REPORT, *supra* note 1, at 32–33 (“Casework is not scientifically valid research, and experience alone cannot establish scientific validity.”).

⁶⁶ Letter from Michael A. Ramos to President Barack Obama, *supra* note 65, at 8.

ommendations related to the admissibility of forensic science evidence.”⁶⁷

More recently, President Trump’s appointment for Attorney General, Jeff Sessions, took further steps to halt forensic science reform efforts. For instance, Sessions recently announced he would halt a collaboration on forensic science research between leading scientists and the Department of Justice (DOJ).⁶⁸ In 2013, responding to the NRC Report, the Obama administration formed the National Commission on Forensic Science (NCFS), a panel of leading scientists charged with advising the executive branch on efforts to standardize and validate forensic disciplines.⁶⁹ But in April of 2017, Sessions announced that the NCFS charter would not be renewed, over the objection of several of NCFS’s scientists.⁷⁰ “For too long,” the panel members wrote, “decisions regarding forensic science have been made without the input of the research science community.”⁷¹ This exclusion, they argued, led to a “disconnection between the fundamental principles of science and some forensic disciplines.”⁷²

In addition to allowing the NCFS charter to expire, Sessions also announced that he would be terminating a review of FBI testimony in several forensic disciplines. Reviews in recent years had already turned up widespread problems with the testimony of FBI analysts in the area of microscopic hair analysis, and Deputy Attorney General Sally Yates had announced that expanded review would focus on identifying “whether the same kind of ‘testimonial overstatement’ that we found during our review of microscopic hair evidence could have crept into other disciplines.”⁷³ This review has now been “suspended”

⁶⁷ Gary Fields, *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. (Sept. 20, 2016, 4:25 PM), <https://www.wsj.com/articles/white-house-advisory-council-releases-report-critical-of-forensics-used-in-criminal-trials-1474394743> (quoting statement of U.S. Attorney General Loretta Lynch).

⁶⁸ Spencer S. Hsu, *Sessions Orders Justice Dept. to End Forensic Science Commission, Suspend Review Policy*, WASH. POST (Apr. 10, 2017), https://www.washingtonpost.com/local/public-safety/sessions-orders-justice-dept-to-end-forensic-science-commission-suspend-review-policy/2017/04/10/2dada0ca-1c96-11e7-9887-1a5314b56a08_story.html.

⁶⁹ *See id.*

⁷⁰ *Id.*

⁷¹ Letter from Comm’rs, NCFS, to Hon. Jefferson B. Sessions, U.S. Attorney Gen. & Dr. Kent Rochford, Acting Dir. of Nat’l Inst. of Standards and Tech. (Apr. 6, 2017), <https://assets.documentcloud.org/documents/3549346/Scientists-on-national-commission-urge-panel-be.pdf>.

⁷² *Id.*

⁷³ Sally Q. Yates, Deputy Attorney Gen., Remarks at the Ninth Meeting of the National Commission on Forensic Science (Mar. 21, 2016), <https://www.justice.gov/opa/speech/deputy-attorney-general-sally-q-yates-delivers-remarks-ninth-meeting-national-commission>.

by the Trump administration.⁷⁴ As Peter Neufeld, co-founder of the Innocence Project, said of these developments, “the [DOJ] has literally decided to suspend the search for the truth.”⁷⁵

As these actions on the part of the DOJ demonstrate, prosecutors’ influence extends beyond the courtroom and into policy and funding decisions. The fact that Sessions could unilaterally disband the NCFS, for example, is a product of the DOJ’s efforts under Lynch to maintain control of forensic science reform. It would therefore be shortsighted to blame the Trump administration for the forensic science problem, since in the months following publication of the PCAST Report in September of Obama’s final year in office, “not only [did] the Obama administration [do] nothing about the report, the Justice Department . . . publicly denounced it.”⁷⁶ As long as forensic science continues to be “powerful evidence in a court of law,” as the NDAA put it,⁷⁷ the DOJ is likely to exercise that influence to keep reform efforts within their control. And if courts give the benefit of the doubt to prosecutors as to the reliability of their evidence, prosecutors will have no incentive to fund research that might prove otherwise.

In sum, neither the structural incentives for prosecutors nor their actual practices suggest that prosecutors are likely to solve the forensic science problem.

B. Criminal Defense Attorneys

Criminal defense attorneys, unlike prosecutors, are structurally incentivized to challenge forensic testimony, yet commentators have noted that they rarely do so.⁷⁸ There are a few possible explanations for this phenomenon. First, because most criminal defendants are indigent, most criminal defense attorneys are public defenders or panel attorneys.⁷⁹ Average caseload numbers for public defenders far

⁷⁴ Hsu, *supra* note 68.

⁷⁵ *Id.*

⁷⁶ Radley Balko, Opinion, *When Obama Wouldn't Fight for Science*, WASH. POST (Jan. 4, 2017), https://www.washingtonpost.com/news/the-watch/wp/2017/01/04/when-obama-wouldnt-fight-for-science/?utm_term=.2e6515f74e0c.

⁷⁷ Letter from Michael A. Ramos to President Barack Obama, *supra* note 65, at 8.

⁷⁸ See, e.g., Neufeld, *supra* note 14, at S110 (noting that although there were nearly fifteen million criminal charges filed in state courts in the year 2000, a review of published opinions from state criminal courts from August 1999 through August 2000 turned up only fifty defense challenges to admissibility).

⁷⁹ Alexa Van Brunt, *Poor People Rely on Public Defenders Who Are Too Overworked to Defend Them*, GUARDIAN (June 17, 2015, 7:30 AM), <https://www.theguardian.com/commentisfree/2015/jun/17/poor-rely-public-defenders-too-overworked> (noting that more than eighty percent of people facing felony charges are indigent and rely on representation by a public defender); see also CAROLINE WOLF HARLOW, BUREAU OF JUSTICE

exceed recommended numbers,⁸⁰ forcing them to cut corners and file only the most critical motions, which could explain the relative quiet when it comes to motions regarding forensic science evidence.⁸¹ Panel attorneys are seldom compensated at competitive rates for their public appointments,⁸² incentivizing them to spend as little time as possible on their indigent clients relative to their paying ones. They also often lack experience with criminal trials.⁸³ And given the frequency with which forensic evidence is admitted,⁸⁴ even defense attorneys with time and money to burn might be reluctant to occupy the court's time with a motion challenging a prosecutor's proffer, fearing that doing so will annoy trial judges who have been admitting such evidence for years. That said, criminal defense attorneys do challenge experts sometimes, and will likely do so more often in light of the growing chorus of criticisms from the NRC Report, the PCAST

STATISTICS, DEFENSE COUNSEL IN CRIMINAL CASES 1 (2000), <https://www.bjs.gov/content/pub/pdf/dccc.pdf> (“At felony case termination, court-appointed counsel represented 82% of State defendants in the 75 largest counties in 1996 and 66% of Federal defendants in 1998.”).

⁸⁰ See Jaeah Lee, Hannah Levintova & Brett Brownell, *Charts: Why You're in Deep Trouble if You Can't Afford a Lawyer*, MOTHER JONES (May 6, 2013, 10:00 AM), <http://www.motherjones.com/politics/2013/05/public-defenders-gideon-supreme-court-charts> (comparing the recommended caseload numbers to the reality across cases of different levels of complexity and concluding that public defenders are overburdened).

⁸¹ See Neufeld, *supra* note 14, at S110 (“[E]ven if inclined to mount a *Daubert* challenge, [criminal defense attorneys] lack the requisite knowledge and skills, as well as the funds, to succeed.”).

⁸² See Peter A. Joy & Kevin C. McMunigal, *Does the Lawyer Make a Difference? Public Defender v. Appointed Counsel*, 27 CRIM. JUST. 46, 47 (2012) (describing results of a Rand study which found that court-appointed lawyers in Philadelphia earn about two dollars per hour, which “fail[s] to attract qualified lawyers, discourage[s] adequate preparation, and create[s] an incentive for appointed lawyers to take on many more cases than they can adequately handle”).

⁸³ See, e.g., *Does Equal Justice for All Include the Poor?*, NPR (Nov. 5, 2013, 11:49 AM), <http://www.npr.org/templates/story/story.php?storyId=243213638> (quoting Professor Eve Primus discussing various systems of indigent defense, including appointment of counsel who “may or may not have any experience in criminal law”). Despite the overwhelming caseloads that burden public defenders, studies have shown that outcomes are better for defendants with public defenders than those who receive appointed counsel. See, e.g., Joy & McMunigal, *supra* note 82, at 46 (describing a Rand study of the Philadelphia area, which found that public defender representation compared to appointed counsel representation reduced a murder defendant's conviction rate by nineteen percent, reduced the likelihood of a life sentence by sixty-two percent, and decreased the expected prison term by twenty-four percent).

⁸⁴ See Neufeld, *supra* note 14, at S110 (“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.”).

Report, and other commentators. Their motions, however, are routinely denied.⁸⁵

Still, when bad forensic evidence is admitted—whether in the absence of a motion to exclude, or in the face of one—defense attorneys can challenge this evidence with the usual tricks of the trade: rigorous cross-examination and dueling experts.⁸⁶ But neither of these traditional safeguards is sufficient. Cross-examination, for example, comes only after the jury has watched the judge—ostensibly the only “neutral” lawyer in the room—qualify the expert at the outset of his or her testimony.⁸⁷ Jurors likely attribute at least a modicum of reliability to someone whose credentials have just been read and used as the basis for admittance.⁸⁸ Effective cross-examination also requires the defense attorney to be relatively knowledgeable about the shortcomings of the evidence, which may be too tall an order for overwhelmed public defenders or appointed counsel more accustomed to writing wills or appearing in bankruptcy court.⁸⁹ A study of 137 transcripts from the trials of DNA exonerees exposed routine failure on the part of defense attorneys to effectively cross-examine expert witnesses, even when their testimonies were patently false.⁹⁰ And finally, even a great cross-examination may not be effective against a confident expert witness, since cross-examination is best suited to exposing personal flaws like “[v]eracity, memory, motivation, prejudices, and

⁸⁵ See *infra* notes 142–45 and accompanying text (describing studies finding that judges reject defense proffers at substantially greater rates than proffers from prosecutors or civil parties).

⁸⁶ See *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 596 (1993) (“Vigorous 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.”).

⁸⁷ See Caroline T. Parrott et al., *Differences in Expert Witness Knowledge: Do Mock Jurors Notice and Does It Matter?*, 43 J. AM. ACAD. PSYCHIATRY & L. 69, 77 (2015) (hypothesizing that jurors may be influenced by “the peripheral cue” that “being an expert extend[s] a blanket influence of knowledge,” and may assume that the judge would “allow[] only qualified people with specialized knowledge to take the role of expert”); see also FED. R. EVID. 702 advisory committee’s note to 2000 amendments (advising courts not to use the word “expert” in order to avoid “put[ting] their stamp of authority” on the testimony) (quoting Charles R. Richey, *Proposals to Eliminate the Prejudicial Effect of the Use of the Word “Expert” Under the Federal Rules of Evidence in Civil and Criminal Jury Trials*, 154 F.R.D. 537, 559 (1994)).

⁸⁸ See *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 . . .”).

⁸⁹ For an interesting account of an appointed lawyer with no criminal experience, see *Deep End of the Pool*, THIS AMERICAN LIFE (Aug. 28, 2016), <https://www.thisamericanlife.org/radio-archives/episode/595/transcript>.

⁹⁰ Garrett & Neufeld, *supra* note 3, at 89.

biases,” rather than scientific invalidity.⁹¹ Cross-examination is therefore both rare and likely ineffective as a tool to prevent invalid forensic science from swaying a jury toward conviction.

In light of these shortcomings of cross-examination, one could argue that contrary experts proffered by the defense are the best safeguard to prevent the jury from crediting the invalid testimony of a prosecution expert. Defense attorneys experience problems in this arena as well, however. First, they may not have access to qualified experts, either for lack of funding⁹² or because the majority of experts are employed by the state.⁹³ Judges may also admit the state’s expert and then turn around and bar a similarly credentialed expert from the defense: At least one study has shown that judges are significantly more likely to exclude experts proffered by the defense than those proffered by the state.⁹⁴ And even if they are admitted, juries faced with conflicting expert testimonies may discount the testimonies of *both* experts, perhaps interpreting them as “hired guns” who are simply testifying as directed.⁹⁵ Thus the traditional safeguards of the trial process are often too little, too late.

⁹¹ Bert Black et al., *Science and the Law in the Wake of Daubert: A New Search for Scientific Knowledge*, 72 TEX. L. REV. 715, 789 (1994) (noting that “most commentators believe ostensibly scientific testimony may sway a jury even when as science it is palpably wrong”).

⁹² See ABA STANDING COMM. ON LEGAL AID & INDIGENT DEFENDANTS, *Gideon’s Broken Promise: America’s Continuing Quest for Equal Justice* 38 (2004) (finding that funding for indigent defense counsel is “shamefully inadequate,” leaving attorneys without “bare necessities for an adequate defense” including experts); Garrett & Neufeld, *supra* note 3, at 89–90 (“[C]ourts frequently deny the defense funding for experts in criminal cases in which forensic evidence plays a central role.”); Neufeld, *supra* note 14, at S110 (“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.”).

⁹³ See Paul C. Giannelli, *Ake v. Oklahoma: The Right to Expert Assistance in a Post-Daubert, Post-DNA World*, 89 CORNELL L. REV. 1305, 1327, 1331 (2004) (noting that while prosecutors have access to “state, county, regional, or metropolitan crime laboratories,” more than half of these laboratories will not test evidence for defense attorneys, giving prosecutors an “overwhelming advantage” in securing expert assistance).

⁹⁴ Groscup & Penrod, *supra* note 15, at 1155, 1165 tbl.1 (in a study of more than 1800 cases, finding that judges admitted 88.6% of prosecution experts and 24.4% of defense experts).

⁹⁵ Lora M. Levett & Margaret Bull Kovera, *Psychological Mediators of the Effects of Opposing Expert Testimony on Juror Decisions*, 15 PSYCHOL. PUB. POL’Y & L. 124, 127 (2009). The “hired gun” heuristic may be applied more often to defense experts than to those proffered by the state, since defense counsel are more likely to rely on “experts for hire” rather than state employees whose compensation is not directly linked to trial testimony. Studies have suggested that, particularly where the testimony in question is complex, juries may rely on peripheral cues, including how often experts testify and how much they are compensated, to weigh the value of their testimony, rather than the content of the testimony itself. See, e.g., Joel Cooper & Isaac M. Neuhaus, *The “Hired Gun” Effect: Assessing the Effect of Pay, Frequency of Testifying, and Credentials on the Perception of*

C. Judges

Although prosecutors lack structural incentives to police forensic evidence appropriately, and defense attorneys lack the resources and authority to do so, judges do not face either of these obstacles. As the neutral arbiter of the courtroom, the judge is tasked with ensuring the fairness and integrity of criminal proceedings. This nebulous expectation is concretized in the trial judge's affirmative obligation to screen all expert evidence before it is presented at trial, known as the judicial gatekeeping function.

All fifty states and the federal government are bound by rules of evidence that require judges to screen expert evidence before it is presented to a jury.⁹⁶ For example, Federal Rule of Evidence 702 states that “[a] witness who is qualified as an expert . . . may testify in the form of an opinion or otherwise” only if certain criteria are met, including that “the testimony is based on sufficient facts or data,” that “the testimony is the product of reliable principles and methods,” and that those principles and methods were “reliably applied” in the case at bar.⁹⁷ Although there's no question that judges benefit from rigorous challenges to evidence by criminal defense attorneys,⁹⁸ the fact remains that with or without a defense motion, judges are obligated to protect the jury from certain expert evidence.⁹⁹

The majority of states and the federal government apply the admissibility test established by the Supreme Court in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*,¹⁰⁰ known as the *Daubert* test.¹⁰¹ *Daubert* superseded the longstanding test articulated in 1923 by the Court of Appeals of the District of Columbia in *Frye v. United States*.¹⁰² The *Frye* standard asks simply whether the methods or principles in question “have gained general acceptance in the field in

Expert Testimony, 24 LAW & HUM. BEHAV. 149, 162 (2000) (finding that mock jurors perceived highly compensated experts who testified frequently to be “less likeable, honest, trustworthy, [and] believable”).

⁹⁶ Shelton, *Judicial Bias*, *supra* note 20, at 22–23 (providing the results of an ABA survey of the expert evidence tests in all fifty states).

⁹⁷ FED. R. EVID. 702; *see also* FED. R. EVID. 702 advisory committee's note to 2000 amendments (noting that Rule 702 was amended after *Daubert* to “affirm[] the trial court's role as gatekeeper”).

⁹⁸ In fact, at least one federal judge has granted a habeas petition on an ineffective assistance of counsel theory because a criminal defense attorney failed to make a *Daubert* motion. *See United States v. Hebshie*, 754 F. Supp. 2d 89, 120–22 (D. Mass. 2010).

⁹⁹ Because of this obligation, the forensic science problem cannot exclusively—or even primarily—be attributed to non-judicial actors.

¹⁰⁰ 509 U.S. 579 (1993).

¹⁰¹ Shelton, *Judicial Bias*, *supra* note 20, at 22–23 (showing that thirty-five states apply the *Daubert* test in at least some circumstances).

¹⁰² 293 F. 1013 (D.C. Cir. 1923).

which it belongs,” known as the “general acceptance” test.¹⁰³ In *Daubert*, the Court interpreted Rule 702 to mandate a judicial finding of “evidentiary reliability” for admittance.¹⁰⁴ The *Daubert* opinion provides four factors for courts to consider in determining the reliability of expert evidence, including (1) whether the theory or technique at issue can be tested, (2) whether it has been subject to peer review and publication, (3) the known rate of error, and (4) whether the theory or technique is generally accepted in the relevant scientific community.¹⁰⁵ The Court emphasized, however, that the test is “flexible”; these factors are neither mandatory nor are they necessarily exhaustive considerations.¹⁰⁶ The Court has subsequently written that the *Daubert* factors “do not constitute a ‘definitive checklist or test’”¹⁰⁷ and their applicability “depend[s] on the nature of the issue, the expert’s particular expertise, and the subject of his testimony.”¹⁰⁸

Because *Daubert* interprets Rule 702, which applies identically to civil and criminal cases, the case requires judges to apply the same “exacting standard[]”¹⁰⁹ to expert proffers from any party. But the flexibility of the *Daubert* test makes it ripe for biased decisionmaking. Over the course of three cases, known as the “*Daubert* trilogy,”¹¹⁰ the Supreme Court granted trial judges discretion on three levels. First, it is up to the judge to decide what kind of procedure to follow when making an admissibility determination—for instance, whether to hold

¹⁰³ *Id.* at 1014. A sizeable minority of states continue to follow the *Frye* test for some or all of their evidentiary decisions, while a few states, including Utah, Virginia, and Wisconsin, use neither *Daubert* nor *Frye*. Shelton, *Judicial Bias*, *supra* note 20, at 22–23. As the dominant rule, *Daubert* is the focus of this Note. However, it is worth noting that much of forensic science ought to be barred under *Frye*’s “general acceptance” standard. The most obvious example of forensic evidence that falls short of this standard is what has been described as “maverick” forensics, evidence offered by experts who “claim to have techniques that have not been heard of or tested, that are based on nothing but the witness’s say-so, that others in the field believe do not exist, and that can be performed only by the maverick.” Moriarty & Saks, *supra* note 18, at 29.

¹⁰⁴ *Daubert*, 509 U.S. at 590.

¹⁰⁵ *Id.* at 593–94. In *Kumho Tire Co. v. Carmichael*, the Supreme Court made clear that *Daubert*’s gatekeeping obligation applied to all forms of Rule 702 evidence, not only scientific evidence. 526 U.S. 137, 147 (1999).

¹⁰⁶ *Daubert*, 509 U.S. at 594 & n.12.

¹⁰⁷ *Kumho Tire*, 526 U.S. at 150 (quoting *Daubert*, 509 U.S. at 593).

¹⁰⁸ *Id.* (quoting Brief for United States as Amicus Curiae Supporting Petitioners at 19, *Kumho Tire*, 526 U.S. 137 (No. 97-1709)).

¹⁰⁹ *Weisgram v. Marley Co.*, 528 U.S. 440, 455 (2000) (noting that *Daubert* put parties on notice of the standards required for admissibility of evidence).

¹¹⁰ The *Daubert* trilogy includes: *Daubert*, 509 U.S. 579, *General Electric Co. v. Joiner*, 522 U.S. 136, 143 (1997) (establishing that admissibility rulings are reviewed for abuse of discretion only), and *Kumho Tire*, 526 U.S. 137 (extending *Daubert*’s holding to non-scientific experts).

a *Daubert* hearing as opposed to deciding the issue on paper.¹¹¹ Second, which, if any, *Daubert* factors to consider is reserved to the judge's discretion.¹¹² And third, the ultimate decision of whether to admit the evidence is a matter of discretion.¹¹³ Recognizing the multiple layers of deference developed through the *Daubert* doctrine, Justice Scalia concurred in one of the trilogy cases with the sole purpose of cautioning that the discretion that trial judges enjoy "is not discretion to perform the function inadequately" but is only "discretion to choose among *reasonable* means of excluding expertise that is *fausse* and science that is junky."¹¹⁴

In the aftermath of the *Daubert* trilogy, scholars predicted a sea change in the admission of forensic science.¹¹⁵ But within a few years, those predictions were replaced with observations that only prosecutors' experts seemed to be evading review. A famous (or infamous) pair of rulings from the Eastern District of Pennsylvania in *United States v. Llera Plaza*¹¹⁶ illustrates both how *Daubert* can be applied to bar shoddy forensic evidence, and how forensic evidence makes it into court despite its shortcomings.

In *Llera Plaza I*, Judge Louis Pollack applied the relatively new *Daubert* standard to fingerprint evidence proffered by the government, and found it lacking in all four factors. First, as to whether the technique "can be (and has been) tested,"¹¹⁷ Judge Pollack found that the methodology used by the FBI had not been subject to relevant testing, rejecting the government's contention that "testing" included

¹¹¹ See *Kumho Tire*, 526 U.S. at 152 ("The trial court must have . . . latitude in deciding how to test an expert's reliability . . . and to decide whether or when special briefing or other proceedings are needed to investigate reliability . . .").

¹¹² See *id.* at 150.

¹¹³ See *Joiner*, 522 U.S. at 142.

¹¹⁴ *Kumho Tire*, 526 U.S. at 159 (Scalia, J., concurring).

¹¹⁵ See, e.g., Murphy, *supra* note 9, at 621 (noting that *Daubert* was "heralded as a watershed moment in the treatment of scientific evidence" that many commentators believed could "stem the perceived epidemic of 'junk science' that had overtaken the courts"); Michael J. Saks & Jonathan J. Koehler, *The Coming Paradigm Shift in Forensic Identification Science*, 309 SCIENCE 892, 894–95 (2005) (citing *Daubert* as one reason for their predicted paradigm shift); see also Erica Beecher-Monas, *Blinded by Science: How Judges Avoid the Science in Scientific Evidence*, 71 TEMP. L. REV. 55, 57 (1998) (arguing that the *Daubert* decision presented an opportunity for courts to begin screening out junk forensic evidence).

¹¹⁶ *United States v. Llera Plaza (Llera Plaza I)*, 179 F. Supp. 2d 492 (E.D. Pa. 2002), vacated 188 F. Supp. 2d 549 (E.D. Pa. 2002); *United States v. Llera Plaza (Llera Plaza II)*, 188 F. Supp. 2d 549 (E.D. Pa. 2002).

¹¹⁷ *Llera Plaza I*, 179 F. Supp. 2d at 503 (quoting *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 593 (1993)).

“‘adversarial’ courtroom testing.”¹¹⁸ Second, he considered whether fingerprinting had been subject to peer review, and concluded that because the only “peer review” came from the analysts themselves, who “tend to be skilled professionals who have learned their craft on the job and without concomitant advanced academic training,”¹¹⁹ the FBI’s methodology had not been subject to “peer review” in the scientific sense.¹²⁰ Third, Judge Pollack found evidence regarding the error rate of fingerprint identification “unpersuasive, one way or another” and found that fingerprint comparison lacked “uniformly accepted ‘scientific’ standards.”¹²¹ Fourth, he found that because the field of fingerprint analysis is not itself a reliable field, general acceptance within the field of fingerprint examiners “by itself cannot sustain the government’s burden in making the case for admissibility of fingerprint testimony under [Rule] 702.”¹²² In sum, fingerprint analysis had not been tested, was not standardized, could not support claims of a low error rate, and was not generally accepted within a disinterested scientific community.¹²³

Nevertheless, after finding fingerprint analysis lacking at every turn, Judge Pollack ruled that fingerprint examiners *could* testify, albeit cabining them to “descriptive, not judgmental” testimony.¹²⁴ In so doing, he alluded to a consideration not sanctioned by the *Daubert* Court, namely the “century of judicial acquiescence in fingerprint identification” that would render total exclusion “unwarrantably heavy-handed.”¹²⁵

Then, the plot thickened. Under protestations from the government that cabining expert testimony would “seriously compromise[]” the government’s “prosecutorial effectiveness,” Judge Pollack held another *Daubert* hearing and overturned his original ruling.¹²⁶ In *Llera Plaza II*, the judge did not dispute the defense’s contention that no new factual or legal basis had emerged to justify the reconsideration, but simply lowered the bar for the FBI. For instance, instead of hewing to his original judgment that the government had not per-

¹¹⁸ *Id.* at 506. Judge Pollack pointed out that “to rely on [the adversary process] would be to vitiate the gatekeeping role of federal trial judges, thereby undermining the essence of Rule 702 as interpreted by the Court in *Daubert*.” *Id.*

¹¹⁹ *Id.* at 509.

¹²⁰ *Id.*

¹²¹ *Id.* at 516.

¹²² *Id.* at 515.

¹²³ *Id.* This critique was largely confirmed by the NRC Report in 2009. NRC REPORT, *supra* note 5, at 142–45.

¹²⁴ *Llera Plaza I*, 179 F. Supp. 2d at 516.

¹²⁵ *Id.*

¹²⁶ *Llera Plaza II*, 188 F. Supp. 2d 549, 552–53, 576 (E.D. Pa. 2002).

suaded him that the error rate was acceptably low,¹²⁷ Judge Pollack now held that “there is no evidence that the error rate of certified FBI fingerprint examiners is unacceptably high.”¹²⁸ This way of looking at the error rate is all the more remarkable because under Rule 702 the burden is clearly on the government to show that the evidence is reliable by a preponderance of the evidence.¹²⁹ Flipping this standard on its head, Judge Pollack concluded that he had not been “persuaded that courts should *defer* admission of testimony with respect to fingerprinting.”¹³⁰ In short, under pressure from the DOJ and the FBI, Judge Pollack put the burden on the defense to show *unreliability* rather than holding the prosecution to its burden of proving *reliability*.

The layers of discretion baked into *Daubert* open the door for this kind of special treatment. For instance, in the Ninth Circuit’s *Daubert* opinion following remand from the Supreme Court, Chief Judge Kozinski offered additional factors for trial judges to consider in making *Daubert* rulings, and called it “very significant” whether the expert’s testimony “grow[s] naturally and directly out of research they have conducted independent of the litigation, or whether they have developed their opinions expressly for purposes of testifying.”¹³¹ And yet, with no explanation for the distinction, he dropped a footnote exempting “law enforcement” evidence from this consideration:

There are, of course, exceptions. Fingerprint analysis, voice recognition, DNA fingerprinting and a variety of other scientific endeavors closely tied to law enforcement may indeed have the courtroom as a principal theatre of operations. As to such disciplines, the fact that the expert has developed an expertise principally for purposes of litigation will *obviously* not be a substantial consideration.¹³²

Interestingly, Judge Kozinski has now joined the crusade against “voodoo science” proffered by prosecutors.¹³³ But his 1995 opinion

¹²⁷ *Llera Plaza I*, 179 F. Supp. 2d at 516 (“[T]he court finds that the information of record is unpersuasive, one way or another, as to [the methodology’s] ‘scientific’ rate of error . . .”).

¹²⁸ *Llera Plaza II*, 188 F. Supp. 2d at 566. Judge Pollack thus concluded despite crediting the defense expert’s criticisms of the FBI’s proficiency tests. “On the record before me, the FBI examiners got very high proficiency grades, but the tests they took did not.” *Id.* at 565.

¹²⁹ FED. R. EVID. 702 advisory committee’s note to 2000 amendments (citing *Bourjaily v. United States*, 483 U.S. 171 (1987)). A contrary rule would have the absurd result of requiring that criminal defendants (the vast majority of whom are indigent) fund proficiency testing for any dicey forensic discipline they hope to keep out of court. Yet *Llera Plaza II* seems to suggest just such a rule. See *Llera Plaza II*, 188 F. Supp. 2d at 566, 572.

¹³⁰ *Llera Plaza II*, 188 F. Supp. 2d at 572 (emphasis added).

¹³¹ *Daubert v. Merrell Dow Pharm., Inc.*, 43 F.3d 1311, 1317 (9th Cir. 1995).

¹³² *Id.* at 1317 n.5 (emphasis added) (internal citations omitted).

¹³³ Alex Kozinski, Opinion, *Rejecting Voodoo Science in the Courtroom*, WALL ST. J. (Sept. 19, 2016, 7:36 PM), <https://www.wsj.com/articles/rejecting-voodoo-science-in-the->

reflects the evidently popular sentiment that law enforcement evidence is somehow uniquely exempt from the rigorous screening that other evidence must pass through.

Prosecutors, criminal defense attorneys, and trial judges are all complicit to some degree in the admission of shoddy forensic science evidence. Only trial judges, however, are explicitly tasked with keeping unreliable expert evidence out of the courtroom. The existence of this affirmative obligation, however, has not prevented unreliable evidence proffered by prosecution from being admitted in court and functioning as the basis for criminal convictions. The following Part attempts to make sense of this reality.

III

UNDERSTANDING THE JUDICIAL FAILURE TO GATEKEEP

As the preceding discussion shows, the forensic science problem is unlikely to be solved by either prosecutors or defense attorneys. That would seem to leave us with trial judges, and indeed Rule 702 and *Daubert* suggest that judges have the means to stop bad forensic evidence in its tracks. But judges and other commentators have suggested a number of reasons why judges do not step in to keep bad forensic expert evidence out of court. This section considers those arguments, and concludes that despite the obstacles in their paths, judges can and should take a leadership role in solving the forensic science problem.

A. *Understanding How Judges Overlook Problems with Forensic Science*

1. *Judges Lack Scientific Knowledge*

One commonly offered explanation for why judges are not well suited to the task of solving the forensic science problem is that they lack scientific training. Indeed, some evidence suggests that judges are not well equipped to operationalize *Daubert* because, though fully capable of reciting the *Daubert* factors, they are not well versed in what those factors actually mean for the evidence in question.¹³⁴

A common problem in the forensic science context, for example, is defining “peer review” for the purposes of *Daubert*’s second

courtroom-1474328199 (“Among the more than 2.2 million inmates in U.S. prisons and jails, countless may have been convicted using unreliable or fabricated forensic science.”).

¹³⁴ See Sophia I. Gatowski et al., *Asking the Gatekeepers: A National Survey of Judges on Judging Expert Evidence in a Post-Daubert World*, 25 LAW & HUM. BEHAV. 433, 444–48 (2001) (discussing results of a survey of 400 state court judges which indicated that only four percent of respondents accurately described the relevance of *Daubert*’s falsifiability and error rate factors).

factor.¹³⁵ In *Daubert* the Court described peer review as “submission to the scrutiny of the scientific community,” which “increases the likelihood that substantive flaws in methodology will be detected.”¹³⁶ One of the most problematic aspects of forensic science is that, because virtually all forensic disciplines (with the notable exception of DNA analysis) were developed for law enforcement purposes, there is no neutral “scientific community” reviewing the methodologies developed in the field. Although some judges have recognized that review by practitioners working in law enforcement cannot reasonably be equated with the kind of dispassionate review described in *Daubert*,¹³⁷ others accept arguments from prosecutors and practitioners that publications and reviews by fellow practitioners satisfy the peer review consideration.¹³⁸

The *Llera Plaza* opinions, analyzed above,¹³⁹ illustrate one type of interpretive error with regard to the peer review factor. In *Llera Plaza I*, Judge Pollack found that fingerprint analysis came up short on peer review because, to the extent that publications were submitted to a peer community, that community was comprised of “skilled professionals who have learned their craft on the job and without any concomitant advanced academic training” and thus was not a “scientific community.”¹⁴⁰ In *Llera Plaza II*, however, Judge Pollack reversed course, this time equating the field of fingerprint analysts to “accountants, vocational experts, accident-reconstruction experts, appraisers of land or of art, experts in tire failure analysis, or others” who “have ‘technical, or other specialized knowledge’” and need not represent a “scientific community” in order to satisfy the “peer review” factor.¹⁴¹ In fact, in both instances Judge Pollack’s analysis overlooks a critical component of peer review, which is that peer review is designed in part to control for bias, and cannot serve that

¹³⁵ See *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 593–94 (1993).

¹³⁶ *Id.* at 593.

¹³⁷ See, e.g., *United States v. Green*, 405 F. Supp. 2d 104, 109 n.7 (D. Mass. 2005) (noting that, while the Association of Firearm and Tool Mark Examiners (AFTE) “publishes a journal that is peer-reviewed by other members of the field,” that “‘field’ consists entirely of individuals who work for law enforcement agencies” as opposed to the field of DNA analysis, which includes “neutral academics as well as law enforcement personnel”).

¹³⁸ See, e.g., *United States v. Otero*, 849 F. Supp. 2d 425, 433 (D.N.J. 2012) (“AFTE theory is subject to peer review through submission to and publication by the AFTE Journal of validation studies which test the theory.”).

¹³⁹ See *supra* notes 116–30 and accompanying text (analyzing the *Llera Plaza* rulings).

¹⁴⁰ *Llera Plaza I*, 179 F. Supp. 2d 492, 509 (E.D. Pa. 2002).

¹⁴¹ *Llera Plaza II*, 188 F. Supp. 2d 549, 563–64 (E.D. Pa. 2002) (“The fact that fingerprint specialists are not ‘scientists,’ and hence that the forensic journals in which their writings on fingerprint identification appear are not ‘scientific’ journals in *Daubert*’s peer review sense, does not seem to me to militate against the utility of the identification procedures employed by fingerprint specialists . . .”).

function adequately when the entire peer group shares a systemic bias. Because fingerprint analysis was developed by law enforcement for law enforcement, the entire peer community may share a systemic bias toward conviction and ought to be approached with skepticism. Having interpreted the Supreme Court's analysis in *Daubert* to mean simply that the peer community ought to be comprised of *highly educated* experts, rather than *neutral* experts, it is no surprise that Judge Pollack backtracked from his initial ruling when confronted with the numerous fields in which scientific training is irrelevant.

But while judges may generally be ill equipped to rigorously evaluate expert evidence proffered by any party, this lack of scientific expertise may unfairly advantage prosecutors and disadvantage defendants. For example, a survey of federal cases in the seven years following the *Daubert* decision found that when criminal defense attorneys challenged prosecution experts, the government prevailed in sixty-one out of the sixty-seven appellate opinions.¹⁴² When the prosecutor was challenging the ruling, on the other hand, the defendant lost forty-four out of fifty-four cases, while seven of the remaining ten cases were remanded for a *Daubert* hearing.¹⁴³ Similar patterns have been identified elsewhere: A subsequent analysis of nearly seven hundred state and federal judicial opinions published in the five-and-a-half years after *Daubert* showed an enormous disparity in success between prosecution and defense proffers of expert evidence at the trial court level, with prosecution experts being admitted 95.8% of the time while defense experts were admitted a mere 7.8% of the time.¹⁴⁴

Prosecution evidence also appears to be held to a lower standard than evidence in civil cases. Empirical studies following *Daubert* have shown that "evidence proffered by plaintiffs in civil cases receives harsh scrutiny for reliability, whereas evidence proffered by prosecu-

¹⁴² D. Michael Risinger, *Navigating Expert Reliability: Are Criminal Standards of Certainty Being Left on the Dock?*, 64 ALB. L. REV. 99, 105 (2000) (noting that only one of the remaining six cases resulted in a reversal based on the unreliability of expert testimony); Neufeld, *supra* note 14, at S109 (analyzing the results of the Risinger study).

¹⁴³ Risinger, *supra* note 142, at 106–07. At the trial court level, in the twelve published opinions addressing defense challenges to prosecution evidence, the evidence was fully admitted in eleven out of the twelve cases and admitted with restrictions in the remaining one. *Id.* at 109. Of the forty-two cases dealing with prosecution challenges to defense proffers, on the other hand, the prosecution prevailed in twenty-eight, or precisely two-thirds of the time. *Id.* at 110.

¹⁴⁴ 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, 346 (2002) (concluding that "[t]he party for whom the key expert testified was significantly related to admission" at trial and on appeal).

tors in criminal cases typically gets a free pass.”¹⁴⁵ Nor is it the case that evidence proffered in civil cases is somehow easier for people without scientific training to understand. As Professor Jane Moriarty wryly puts it, “[i]n civil cases, courts seem quite up to the task of evaluating microbiology, teratology, and toxicology evidence,” but “when it comes to evaluating the shortcomings of lip prints and handwriting, courts are unable to muster the most minimal grasp of why a standardless form of comparison might lack evidentiary reliability.”¹⁴⁶

2. *Judges Do Not Receive Helpful Information from Defense Attorneys*

Judges faced with prosecution expert proffers may not receive the information they need to make a fair gatekeeping determination from the usual source: the opposing party.¹⁴⁷ Ordinarily, judges rely heavily on the arguments and evidence that the parties present to them.¹⁴⁸ As analyzed above, however, defense attorneys frequently fail to challenge prosecution experts.¹⁴⁹ Without those challenges, judges make admissibility determinations solely on the basis of the prosecutor’s expert notice and the judge’s knowledge and beliefs about forensic science.¹⁵⁰ And even where challenges are made, they may be made

¹⁴⁵ Murphy, *supra* note 9, at 621–24 (surveying the empirical studies supporting a civil/criminal divide).

¹⁴⁶ Jane Campbell Moriarty, *Will History Be Servitude?: The NAS Report on Forensic Science and the Role of the Judiciary*, 2010 UTAH L. REV. 299, 315; see also Murphy, *supra* note 9, at 621.

¹⁴⁷ See, e.g., Neufeld, *supra* note 14, at S110 (“Unlike the extremely well-litigated civil challenges, the criminal defendant’s challenge is usually perfunctory. . . . Defense lawyers generally fail to build a challenge with appropriate witnesses and new data.”); Risinger, *supra* note 142, at 135 (finding a “systematic failure to seriously litigate [the weakest kinds of common prosecution-proffered expertise] on the part of the criminal defense bar”). Under those circumstances, judges may be more likely to defer to the prosecutor, who is expected to be more objective than defense counsel. See *supra* note 48 and accompanying text (discussing the prosecutor’s ostensible role as a “minister of justice”).

¹⁴⁸ See David S. Clark, *The Organization of Lawyers and Judges*, in 16 INTERNATIONAL ENCYCLOPEDIA OF COMPARATIVE LAW 3-74 (Mauro Cappelletti ed., 2014) (noting that while judges in civil law systems typically bear the burden for independently researching the law, “the common law system of procedure puts a heavy burden on attorneys to discover the facts in a case and even to brief legal issues for the judge”).

¹⁴⁹ See *supra* notes 78–85 and accompanying text (discussing why defense attorneys seldom file motions challenging prosecution proffers).

¹⁵⁰ It is likely that judges, like most people, generally believe that forensic science is valid and reliable. See Gretchen Gavett, *Judge Harry T. Edwards: How Reliable Is Forensic Evidence in Court?*, PBS (Apr. 17, 2012), <http://www.pbs.org/wgbh/frontline/article/judge-harry-t-edwards-how-reliable-is-forensic-evidence-in-court/> (quoting D.C. Circuit Judge Harry Edwards’s statement that he and his colleagues “assumed that the methodology [of forensic science] was valid and reliable and that the work in putting the evidence together and in offering the testimony was proper” until the NRC uncovered “systemic, serious problems with respect to certain of the [forensic] disciplines”).

without the input—and accompanying affidavit—of a contrary expert witness, and may as a result be substantively weak or lacking in credibility (particularly in contrast to the proffered prosecution expert).¹⁵¹ A judge could reasonably interpret a defense objection unsupported by an expert’s advice and affidavit as a Hail Mary motivated by vigorous advocacy rather than a well-founded challenge to invalid evidence.

Still, with or without a credible defense challenge, judges have a responsibility to keep problematic expert evidence out of the courtroom. As difficult as this may be without the help of effective defense challenges, a growing body of literature—including the NRC Report¹⁵² and the PCAST Report,¹⁵³ as well as scholarship,¹⁵⁴ journalism,¹⁵⁵ and even case law¹⁵⁶—is available to educate judges as they make these decisions. Thus while judges may lack information from the usual source—the opposing party—the information itself is available, and judges have a responsibility to find it.

3. *Judges Rely on Heuristics to Admit Prosecution Evidence*

Judges faced with complex scientific questions for which they have little or no training, often combined with an absence of adequate information from the defense, may fail to grasp the shortcomings of

¹⁵¹ See *supra* notes 92–95 and accompanying text (discussing obstacles defense attorneys experience to obtaining experts, particularly in forensic fields).

¹⁵² NRC REPORT, *supra* note 5 (describing the various types of forensic science, standards of admissibility as evidence in legal cases, and problems in methodology and oversight).

¹⁵³ PCAST REPORT, *supra* note 1 (evaluating the scientific validity of seven types of forensic evidence, and providing specific recommendations for the judiciary).

¹⁵⁴ See *supra* note 3 (collecting sources that critique forensic evidence).

¹⁵⁵ See, e.g., Balko, *supra* note 76 (discussing problems with forensic science and Obama’s failure to take action to solve them); Spencer S. Hsu, *FBI Admits Flaws in Hair Analysis over Decades*, WASH. POST (Apr. 18, 2015), https://www.washingtonpost.com/local/crime/fbi-overstated-forensic-hair-matches-in-nearly-all-criminal-trials-for-decades/2015/04/18/39c8d8c6-e515-11e4-b510-962fcfab310_story.html (discussing the FBI’s announcement that the vast majority of its microscopic hair analysts gave flawed testimony over the two-decade period preceding 2000); Michael Shermer, *Can We Trust Crime Forensics?*, SCI. AM. (Sept. 1, 2015), <https://www.scientificamerican.com/article/can-we-trust-crime-forensics/> (discussing generally problems with forensic science).

¹⁵⁶ For example, in an evidence ruling from the District of New Mexico, Judge William P. Taylor discussed the NRC Report and rulings from other district courts at length before ruling that, while the prosecution’s ballistics expert could testify, he would 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.” *United States v. Taylor*, 663 F. Supp. 2d 1170, 1178, 1180 (D.N.M. 2009). For other cases coming to the same conclusion with regard to ballistics evidence, see *United States v. Glynn*, 578 F. Supp. 2d 567, 570, 574–75 (S.D.N.Y. 2008), *United States v. Diaz*, No. CR 05-00167 WHA, 2007 WL 485967, at *11–14 (N.D. Cal. Feb. 12, 2007), *United States v. Monteiro*, 407 F. Supp. 2d 351, 372–73 (D. Mass. 2006), and *United States v. Green*, 405 F. Supp. 2d 104, 108–09, 114–15 (D. Mass. 2005).

forensic evidence because they are relying on heuristics, or mental shortcuts, to sidestep the substantive question altogether. Humans often rely on heuristics when faced with complex problems, either because they fail to realize that the shortcut is not an appropriate proxy for the more difficult decision, or because they lack the motivation or energy needed to engage in the effortful work demanded.¹⁵⁷

Perhaps the single best example of an inappropriate but tempting heuristic for judges to rely on in the forensic science context is how long certain types of forensic evidence, such as fingerprint analysis, have been admitted as evidence.¹⁵⁸ As one district judge put it, “[T]he methods of latent print identification can be and have been tested. They have been tested for roughly 100 years. They have been tested in adversarial proceedings with the highest possible stakes—liberty and sometimes life.”¹⁵⁹ The problem with this reasoning is that a mistake repeated for a hundred years is still a mistake.¹⁶⁰ Professor Jane Moriarty has illustrated this fallacy by comparing the long-history argument in the forensic science context to similar arguments made to justify medical use of leeches, a practice that survived for thousands of years until it was subjected to scientific testing.¹⁶¹ A long history of admission, in other words, is a poor proxy for validity or reliability of a methodology.

Judges may also rely on their instincts about the parties to guide their decisionmaking when it comes to forensic science experts.¹⁶² Heuristics in this line are particularly concerning because they are likely to militate in favor of the prosecution and against the defendant overall. One reason why such an imbalance may exist is that many

¹⁵⁷ See generally DANIEL KAHNEMAN, THINKING, FAST AND SLOW (2011) (analyzing how and why humans rely on heuristics and biases to facilitate decisionmaking); Christine Jolls & Cass R. Sunstein, *The Law of Implicit Bias*, 94 CALIF. L. REV. 969 (2006) (analyzing the way humans rely on heuristics in the context of anti-discrimination law).

¹⁵⁸ See generally Moriarty, *supra* note 146, at 310–11, 319–20 (arguing that judges unreasonably equate a long history of admission of evidence with reliability).

¹⁵⁹ United States v. Havvard, 117 F. Supp. 2d 848, 854 (S.D. Ind. 2000).

¹⁶⁰ See Moriarty, *supra* note 146, at 316 (“[A] long history of use confers no particular proof of validity.”); see also Edwards & Mnookin, *supra* note 2 (“Respectfully, experience has shown that, at least in criminal trials, the suggestion that the ‘adversarial system’ represents an adequate means of demonstrating the unreliability of forensic evidence is mostly fanciful.”).

¹⁶¹ Moriarty, *supra* note 146, at 316.

¹⁶² For example, cognitive scientists have recognized a heuristic known as the “halo effect.” KAHNEMAN, *supra* note 157, at 82–85. The halo effect describes the human tendency to develop an initial positive or negative impression about someone, and then to fill in knowledge gaps or interpret new information in a way heavily colored by those initial assumptions. *Id.* For example, a professor who grades a single student’s entire exam at once may find themselves heavily influenced in grading the second and third answers by how strong the answer to the first question was. *Id.* at 83–84.

judges were prosecutors before they donned their robes. For example, a review of the judges nominated during President Obama's first seven years in office concluded that while he made significant strides toward diversifying the bench in terms of gender, race, and sexual orientation, the professional experience of his nominees was strikingly homogenous, with the largest number of nominees coming from private practice, followed by prosecutors' offices.¹⁶³ In state courts, the trend is the same. For example, one 2009 study of state supreme court justices found that 32.9% of state supreme court justices had prosecutorial experience, while only 15.4% had experience as public defenders.¹⁶⁴ A high representation of prosecutorial experience on the bench may both reflect and contribute to a judicial preference for prosecutors. And, interlocking with a bias in favor of consistency as described above, former prosecutors-turned-judges may be biased in favor of the evidence itself, as they likely proffered similar forensic evidence during their time as prosecutors—and got convictions because of it.

Furthermore, in some jurisdictions the court system is organized such that prosecutors are assigned to specific courtrooms. Judges and prosecutors assigned to these courtrooms may overlap on a daily basis, leading to a collegial, team-like atmosphere.¹⁶⁵ In the context of heavily discretionary decisionmaking like evidentiary rulings, a friendly relationship with a prosecutor may unconsciously guide a judge's decision to decline a motion for a *Daubert* hearing, discourage her from weighing factors that cut against admission, and ultimately push her toward allowing invalid testimony.

In addition to potential bias in favor of prosecutors, judges may also harbor unconscious biases against criminal defendants. Such biases could come from a number of sources. One factor worth considering is that criminal defendants are disproportionately poor and Black,¹⁶⁶ in a society in which implicit bias against disadvantaged class

¹⁶³ ALLIANCE FOR JUSTICE, BROADENING THE BENCH: PROFESSIONAL DIVERSITY AND JUDICIAL NOMINATIONS 4, 6 (2016), <http://www.afj.org/wp-content/uploads/2014/11/Professional-Diversity-Report.pdf>. Among Obama's circuit court nominees, twenty-four had prosecutorial experience, while only five had experience in public defense. *Id.* at 11. Thirty-one circuit court nominees came from civil government backgrounds. *Id.* At the district court level, 126 nominees were prosecutors, while forty-five were public defenders. *Id.* at 12.

¹⁶⁴ Gregory L. Acquaviva & John D. Castiglione, *Judicial Diversity on State Supreme Courts*, 39 SETON HALL L. REV. 1203, 1235 tbl.10 (2009).

¹⁶⁵ See Walter W. Steele, Jr., *Unethical Prosecutors and Inadequate Discipline*, 38 Sw. L.J. 965, 972 (1984) (expressing concerns about ethical violations under this organizational structure).

¹⁶⁶ See, e.g., MATTHEW R. DUROSE & PATRICK A. LANGAN, BUREAU OF JUSTICE STATISTICS, FELONY SENTENCES IN STATE COURTS, 2002, at 6 (2004) (finding that while

and racial groups “has proven to be extremely widespread.”¹⁶⁷ Given the prevalence of these biases, the judiciary would be in a class by itself if it were unaffected by them.¹⁶⁸ That is not to say that simply because “judges are human”¹⁶⁹ there are no differences in the degree to which they harbor or act on bias. To the contrary, empirical evidence suggests that judges are *not* as susceptible to all cognitive biases as is the population as a whole.¹⁷⁰ That said, studies also indicate that judges are susceptible to cognitive biases, including racial bias.¹⁷¹ It would therefore be too simple to conclude either that judges as a whole are indistinguishable from the general populace, or that, because their jobs demand neutrality, they can simply will away the influence of any spurious considerations.

But although judges may be tempted to rely on heuristics like whether the evidence has a long history of being admitted to court, or instincts about the parties themselves, this is not an insurmountable problem. As the authors of one study of judicial decisionmaking suggest, judges can reduce error and the influence of systemic bias in part by simply making admissibility decisions slowly and deliberately.¹⁷² Similarly, just recognizing that *Daubert* determinations are vulnerable

Black people made up twelve percent of the U.S. population, they comprised thirty-seven percent of the persons convicted of a violent felony); Bernadette Rabuy & Daniel Kopf, *Prisons of Poverty: Uncovering the Pre-Incarceration Incomes of the Imprisoned*, PRISON POLICY INITIATIVE (July 9, 2015), <https://www.prisonpolicy.org/reports/income.html> (finding that incarcerated people were earning forty-one percent less income prior to incarceration than non-incarcerated people of similar ages); see also KAREN DOLAN & JODI L. CARR, INST. FOR POLICY STUDIES, *THE POOR GET PRISON: THE ALARMING SPREAD OF THE CRIMINALIZATION OF POVERTY* (2015) (describing ways in which poverty leads to incarceration and vice versa).

¹⁶⁷ Jolls & Sunstein, *supra* note 157, at 971. See generally John T. Jost et al., *The Existence of Implicit Bias Is Beyond Reasonable Doubt: A Refutation of Ideological and Methodological Objections and Executive Summary of Ten Studies that No Manager Should Ignore*, 29 RES. ORGANIZATIONAL BEHAV. 39 (2009) (describing methods of studying implicit bias and summarizing prominent findings of implicit bias based on race and social status, among other criteria).

¹⁶⁸ See Chris Guthrie et al., *Inside the Judicial Mind*, 86 CORNELL L. REV. 777, 782–83 (2001) (observing that research has exposed cognitive biases among “doctors, real estate appraisers, engineers, accountants, options traders, military leaders, and psychologists”).

¹⁶⁹ Commentators love to point this out. *E.g.*, *id.* at 784 (quoting JEROME FRANK, *COURTS ON TRIAL: MYTH AND REALITY IN AMERICAN JUSTICE* 410 (1949)).

¹⁷⁰ See *id.* at 784, 816–17 (finding in a study of 167 federal magistrate judges that they were just as susceptible as lay people to anchoring, hindsight bias, and egocentric bias, but were less susceptible to framing and the representativeness heuristic).

¹⁷¹ See *id.* at 816–17 (discussing the influence of cognitive illusions on judges); Jeffrey J. Rachlinski et al., *Does Unconscious Racial Bias Affect Trial Judges?*, 84 NOTRE DAME L. REV. 1195, 1205, 1207, 1225 (2009) (finding in a study of 133 judges that implicit racial bias is “widespread”).

¹⁷² Chris Guthrie et al., *Blinking on the Bench: How Judges Decide Cases*, 93 CORNELL L. REV. 1, 36 (2007) (proposing that judges delay making evidentiary rulings until the judge has had time to deliberate).

to heuristic-based decisionmaking may combat the influence of such bias.

B. From Recognition to Action

What the foregoing analysis boils down to is this: Although judges may understandably be hampered in their efforts to recognize the flaws of the forensic evidence proffered by prosecutors, the obstacles to gatekeeping are not so great or so insurmountable as to justify admission of untested or junk evidence. But for judges who recognize the forensic science problem and are prepared to keep it out of court, the final problem may be simply that these judges feel powerless to break with history, particularly where the result is keeping evidence away from the jury.

1. Why Judges May Choose Not to Exclude Bad Evidence

As a threshold matter, it is worth noting that in both state and federal courts, judges are generally encouraged to take a liberal stance with regard to admitting evidence. The baseline requirement for evidence to be admissible is merely that it be “relevant,” meaning that it “has any tendency to make a fact more or less probable than it would be without the evidence,” and that the fact in question “is of consequence in determining the action.”¹⁷³ Emphasizing that this is not a high bar, the comment to Rule 401 advises that “a brick is not a wall,” meaning that evidence need not prove an entire case on its own, it need only be helpful.¹⁷⁴ Relevance, in other words, is a low standard, and forensic evidence easily satisfies it.¹⁷⁵

In addition to the guidance provided by the Federal Rules of Evidence, principles stemming from sources as lofty as the Constitution may encourage a liberal mindset when it comes to the admission of evidence. For example, Judge Jack Weinstein of the Eastern District of New York has provided the following explanation for a liberal approach to evidence rulings:

The jury’s constitutionally based fact-finding primacy demands a measure of forbearance on the part of judges. We cannot forget that, because of our narrow life experiences, our ability to draw appropriate inferences from the evidence in the cases before us is

¹⁷³ FED. R. EVID. 401.

¹⁷⁴ FED. R. EVID. 401 advisory committee’s note on proposed rules (citing KENNETH S. BROUN ET AL., *McCORMICK ON EVIDENCE* § 152 (7th ed. 2016)).

¹⁷⁵ See, e.g., *United States v. Otero*, 849 F. Supp. 2d 425, 438 (D.N.J. 2012) (“[T]hough the parties have not argued it, the relevance of [ballistics] testimony to the charges against Defendants is manifest. Clearly, the evidence will assist the trier of fact to determine a fact in issue . . .”).

limited. Whenever it is arguably appropriate, we should allow the matter to go to the jury, reserving the right to set aside its decision if there proves to be no rational basis for the verdict. Not only is this the fairest approach in most cases, but it also provides litigants with something most desire—a chance to be heard and a judgment by their fairly selected peers.¹⁷⁶

In sum, judges making expert admissibility determinations are encouraged by everything from the rules of evidence to the broadest principles of justice to err on the side of admitting evidence.

In the expert context, there are countervailing principles to consider. As Judge Patti Saris of the District of Massachusetts has opined, “The Court’s vigilant exercise of [the Rule 702] gatekeeper role is critical because of the latitude given to expert witnesses to express their opinions . . . and because an expert’s testimony may be given greater weight by the jury due to the expert’s background and approach.”¹⁷⁷ Yet even where experts are concerned, the Court has implied that trial judges should take a liberal approach to admission, emphasizing that “[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.”¹⁷⁸ A judge who recognizes weaknesses in forensic evidence that is routinely admitted may therefore feel compelled to admit it despite its flaws.

Ironically, the ultimate safeguard against judicial error—appellate review—may actually *discourage* judges from gatekeeping effectively. In at least one jurisdiction, the law is clear that once the admission of a “new scientific technique” has been approved by an appellate court in a written opinion, “the precedent so established may control subsequent trials.”¹⁷⁹ Although many jurisdictions, including federal courts, have not adopted this approach,¹⁸⁰ judges in those jurisdictions are no doubt wary of departing from regular prac-

¹⁷⁶ Jack B. Weinstein, *The Role of Judges in a Government of, by, and for the People: Notes for the Fifty-Eighth Cardozo Lecture*, 30 CARDOZO L. REV. 1, 21 (2008).

¹⁷⁷ *United States v. Monteiro*, 407 F. Supp. 2d 351, 358 (D. Mass. 2006).

¹⁷⁸ *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 596 (1993).

¹⁷⁹ *People v. Kelly*, 549 P.2d 1240, 1245 (Cal. 1976). Even in California, however, this rule comes with the caveat that “new evidence . . . reflecting a change in the attitude of the scientific community” may allow courts to deviate from the precedential decision. *Id.*

¹⁸⁰ *See, e.g., Moore v. Ashland Chem. Inc.*, 151 F.3d 269, 276 (5th Cir. 1998) (describing the *Daubert* standard as one which “requires some objective, independent validation of the expert’s methodology” and demands that the proponent “prove by a preponderance of the evidence that the testimony is reliable”); *Monteiro*, 407 F. Supp. 2d at 356 (rejecting the government’s argument that “because toolmark identification evidence has been deemed admissible by many other courts, the burden of proving such evidence to be unreliable should shift to the defendants”); *Eaton Corp. v. Parker-Hannifin Corp.*, 292 F. Supp. 2d 555, 567 (D. Del. 2003) (“[T]he proponent of the expert testimony . . . must prove by a preponderance of the evidence that the testimony is reliable.”).

tice. Former Judge Nancy Gertner of the District of Massachusetts, an outspoken skeptic of forensic evidence,¹⁸¹ alludes to this concern in *United States v. Green*.¹⁸² “Although the scholarly literature is extraordinarily critical [of ballistics evidence],” she writes, “court after court has continued to allow the admission of this testimony.”¹⁸³ Judge Gertner concludes that precedent unanimously militated in favor of admitting the evidence, even while she described this type of reasoning as “troubling.”¹⁸⁴ “It runs the risk of ‘grandfathering in irrationality,’” she cautions, “without reexamining it in the light of *Kumho* and *Daubert*.”¹⁸⁵

Even if judges believe that exclusion is the correct outcome, they may also be influenced by the threat of political backlash. This concern is likely particularly strong for elected judges. Judicial elections both motivate judicial candidates to espouse “tough on crime” views, and may weed out candidates who are not suitably tough on crime.¹⁸⁶ Defense attorneys who run for judicial seats have been attacked for upholding “the rights of violent criminals,”¹⁸⁷ and incumbent judges have been criticized by opponents for having used “loophole[s]” to

¹⁸¹ See *United States v. Hebshie*, 754 F. Supp. 2d 89, 114–15 (D. Mass. 2010) (suggesting that arson analysis is “bad science” and emphasizing that the admissibility standard for evidence in criminal cases is identical to the civil standard); Gertner, *supra* note 2 (criticizing shoddy forensic evidence and judges who uncritically admit it).

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

¹⁸³ *Id.* at 122.

¹⁸⁴ *Id.* at 123.

¹⁸⁵ *Id.* Judge Gertner admitted the ballistics testimony in *Green*, albeit cabining the expert to “testify[ing] to his observations” without “conclud[ing] that the match he found by dint of the specific methodology he used permits ‘the exclusion of all other guns’ as the source of the shell casings.” *Id.* at 124. If a judge is actually considering departing from the routine practice of her jurisdiction or the judiciary as a whole, there is yet another incentive to toe the party line: the prospect of undermining previous convictions. This concern may be animated by specific cases over which the judge in question presided in the past, or it may be a more general concern about the convictions secured on the basis of a certain type of forensic testimony in general. Thus, a judge faced with Defendant F may have at the back of her mind the past convictions of Defendants A–E when she rules to admit evidence that is increasingly being called into question outside the courtroom.

¹⁸⁶ See, e.g., KATE BERRY, BRENNAN CENTER FOR JUSTICE, HOW JUDICIAL ELECTIONS IMPACT CRIMINAL CASES 4 (2015) (describing one attack ad that targeted former Kentucky Supreme Court Justice Janet Stumbo for having “sided with criminals 50 percent of the time” and noting that Stumbo ultimately lost the election); see also Liz Seaton, *Smear Campaigns Against Justice*, U.S. NEWS (Feb. 29, 2016, 8:00 AM), <https://www.usnews.com/opinion/articles/2016-02-29/supreme-court-sees-how-judicial-elections-mar-criminal-justice-reform> (finding that more than half of the air time for judicial election advertising was focused on tough-on-crime messages).

¹⁸⁷ Joanna Cohn Weiss, Note, *Tough on Crime: How Campaigns for State Judiciary Violate Criminal Defendants’ Due Process Rights*, 81 N.Y.U. L. REV. 1101, 1106 (2006) (quoting Mark Hansen, *When Is Speech Too Free?*, A.B.A. J., May 2001, at 20 (discussing a 2011 California judicial election)).

reduce sentences post-conviction.¹⁸⁸ For example, in the 2014 race for the Arkansas Supreme Court, an ad supporting Judge Robin Wynne touted his “refus[al] to allow technicalities to overturn convictions.”¹⁸⁹ Presumably, *Daubert* reliability is just the kind of “technicality” to which this ad was referring. And judicial elections are more than just a platform for anti-defendant rhetoric: Studies have consistently demonstrated that “the pressures of upcoming re-election and retention election campaigns make judges more punitive toward defendants in criminal cases.”¹⁹⁰ These effects are exacerbated in jurisdictions that have liberal campaign spending laws, and, relatedly, in races with higher levels of TV advertising.¹⁹¹ Although political effects may be felt particularly strongly by elected judges, these findings from the judicial election context suggest that political pressure overall favors prosecutors and disfavors criminal defendants. In other words, to the extent that a trial judge facing an admissibility determination is influenced by political pressure, it is likely to push her toward admitting the evidence.

2. *The Path to Solving the Forensic Science Problem*

The structural limitations that trial judges face in the context of forensic science experts proffered by prosecutors, as well as the legal and policy framework in which they operate, help to explain why judges have not been more aggressive in policing untested or simply invalid forensic science evidence in their courtrooms, despite their affirmative gatekeeping obligation. But while judges may feel hemmed in by the limited information at hand (at least insofar as they are relying on the parties to provide them with information), trial judges, at least in the federal system, have the tools to effectively police shoddy forensic evidence.

Federal Rule of Evidence 702 instructs that a qualified witness may testify as an expert only if “the testimony is based on sufficient facts or data,” “the testimony is the product of reliable principles and

¹⁸⁸ See, e.g., BERRY, *supra* note 186, at 5 (describing an attack ad in which a challenger alleged that the incumbent had released a convicted child rapist who then continued to commit offenses).

¹⁸⁹ *Id.* at 5 (posting a video of *Tough on Predators* (Law Enforcement Alliance of Am. commercial broadcast May 9, 2014), <http://newpoliticsreport.org/spot/leaa-tough-on-predators-2/>); see also Weiss, *supra* note 187, at 1105 (describing an attack campaign from the 1994 race for the Texas Court of Criminal Appeals in which “one candidate and former prosecutor called the Court of Criminal Appeals a ‘citadel of technicality’” (internal citations omitted)).

¹⁹⁰ BERRY, *supra* note 186, at 1 (finding this result consistently across ten “recent, prominent, and widely cited empirical studies” on the subject).

¹⁹¹ *Id.* at 8–9 (discussing the findings of a 2014 American Constitution Society study).

methods,” and “the expert has reliably applied the principles and methods to the facts of the case.”¹⁹² Pursuant to Federal Rule of Evidence 104, the proponent of the evidence bears the burden of showing by a preponderance of the evidence that it meets the requirements of Rule 702.¹⁹³ Absent a sufficient showing, the evidence should be appropriately cabined,¹⁹⁴ or, if no relevant evidence is adequately supported, excluded altogether. Reliability of evidence must be established in each case, and is not a matter of precedent in the federal court system.¹⁹⁵

It may be difficult for judges to stick their necks out and exclude evidence because of an absence of information. Although at least one forensic discipline has been “debunked” by studies affirmatively demonstrating that it is unreliable—namely bite-mark analysis¹⁹⁶—and others have been called into such serious question that the FBI has abandoned them,¹⁹⁷ the majority of the criticisms leveled at forensic disciplines is that they’re untested: We simply don’t know whether they are reliable or not.¹⁹⁸ Under Rule 702, the absence of evidence that a forensic methodology is reliable is a legally sufficient basis for exclusion, but it is nonetheless rhetorically weaker than affirmative

¹⁹² FED. R. EVID. 702.

¹⁹³ See *supra* note 180 (collecting sources describing Rule 104’s requirements).

¹⁹⁴ See *supra* note 156 (collecting cases where testimony was cabined in the context of ballistics evidence).

¹⁹⁵ See *United States v. Monteiro*, 407 F. Supp. 2d 351, 356 (D. Mass. 2006) (rejecting the government’s argument that “because toolmark identification evidence has been deemed admissible by many other courts, the burden of proving such evidence to be unreliable should shift to the defendants”).

¹⁹⁶ See PCAST REPORT, *supra* note 1, at 84, 87 (concluding that additional research into bite-mark analysis is inadvisable because “the prospects of developing bitemark analysis into a scientifically valid method [are] low”).

¹⁹⁷ See, e.g., Microscopic Hair Analysis Press Release, *supra* note 4 (reporting widespread false testimony in the realm of microscopic hair analysis); Press Release, FBI, FBI Laboratory Announces Discontinuation of Bullet Lead Examinations (Sept. 1, 2005), <https://archives.fbi.gov/archives/news/pressrel/press-releases/fbi-laboratory-announces-discontinuation-of-bullet-lead-examinations> [hereinafter CBLA Press Release] (“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.”).

¹⁹⁸ See, e.g., PCAST REPORT, *supra* note 1, at 82 (recommending further study of DNA analysis of complex-mixture samples); *id.* at 96–97 (concluding that “only two properly designed studies of the accuracy of latent fingerprint analysis have been conducted” and recommending further study); *id.* at 111 (finding that “only a single appropriately designed study” of firearms analysis has been conducted, and “[t]here is thus a need for additional, appropriately designed black-box studies to provide estimates of reliability”); *id.* at 117 (concluding that the validity of footwear analysis is “unsupported by any meaningful evidence” and is therefore “not scientifically valid”).

evidence showing, for example, that a methodology has a particularly high rate of error.¹⁹⁹

Nevertheless, trial judges are obligated to ensure that the expert testimony presented in their courtrooms meets the basic standards articulated in Rule 702, and they have the potential to dramatically affect the landscape of forensic science when they exercise that obligation. For example, a forensic discipline known as comparative bullet lead analysis (CBLA) was once commonly admitted in courts.²⁰⁰ CBLA was primarily used in cases where bullets were too mutilated to be subjected to the more common “tool marks” or “ballistics” analysis.²⁰¹ The technique involves analyzing the elements in bullet lead, and applying defined standards to assess whether the lead from the crime scene and the lead from bullets seized from the defendant are “analytically indistinguishable.”²⁰² However, analysts had no basis for the conclusion that, if the bullets were indeed “analytically indistinguishable,” that such a finding suggested they came from the same box of bullets, or even the same geographic region.²⁰³

It was in this context that, in 2003, Judge Ronald Guzmán of the Northern District of Illinois rigorously applied Rule 702 to CBLA, and concluded that the technique came up short.²⁰⁴ In *United States v. Mikos*, Judge Guzmán was confronted with an absence of data to support the claims of the government’s CBLA expert,²⁰⁵ and concluded

¹⁹⁹ Cf. KAHNEMAN, *supra* note 157, at 85–88 (describing the cognitive illusion that “what you see is all there is,” meaning that people are eager to draw conclusions based on the information available and struggle to suspend judgment even when it is clear that they are lacking all relevant information).

²⁰⁰ See Paul C. Giannelli, *Comparative Bullet Lead Analysis: A Retrospective*, 47 CRIM. L. BULL. 306, 306 (2011) (“For over thirty years, FBI experts testified about [CBLA] . . .”).

²⁰¹ See CBLA Press Release, *supra* note 197 (“Bullet lead examinations have historically been performed in limited circumstances, typically when a firearm has not been recovered or when a fired bullet is too mutilated for comparison of physical markings.”).

²⁰² Giannelli, *supra* note 200, at 306.

²⁰³ NAT’L RESEARCH COUNCIL, FORENSIC ANALYSIS: WEIGHING BULLET LEAD EVIDENCE 112–13 (2004) [hereinafter CBLA REPORT] (concluding that because “[d]etailed patterns of distribution of ammunition are unknown . . . an expert should not testify as to the probability that a crime scene bullet came from the defendant”); *id.* at 113 (concluding that because “[t]he available data do not support any statement that a crime bullet came from, or is likely to have come from, a particular box of ammunition” testimony should not assert that “the crime bullet came from the defendant’s box or from a box manufactured at the same time”). Microscopic hair analysis suffers from the same shortcoming, namely that the commonality of hair characteristics in the general population is unknown. See *supra* notes 44–45 and accompanying text.

²⁰⁴ *United States v. Mikos*, No. 02 CR 137, 2003 WL 22922197 (N.D. Ill. Dec. 9, 2003).

²⁰⁵ *Id.* at *1 (“[T]he defense claims that the Government has supplied no underlying basis or support for [the expert’s] conclusion that, because these lead samples are ‘analytically indistinguishable,’ the bullets from the victim and the bullets from the

on that basis that “the required standard of scientific reliability” was met only with regard to the statement that the bullets were “analytically indistinguishable,” and not with regard to the “further opinion that from this finding it follows that the bullets must or even likely came from the same batch or melt.”²⁰⁶ The government’s proffered expert, Charles Peters, contended that one of the elements found in the bullets in question, bismuth, was found in “remarkably unusual” quantities in the two bullets.²⁰⁷ Judge Guzmán was not persuaded by this characterization:

There is . . . absolutely no way to know if the bismuth level in these bullets is in fact “remarkably unusual.” It may be so in Peters’ experience, but as we have pointed out, given the huge population with which we are concerned (in the billions), Mr. Peters’ experience is no more than anecdotal evidence. Such evidence can be particularly misleading because it appears logical and reasonable. If agent Peters has not, in years of experience and after hundreds of analyses, previously encountered such a high quantity of bismuth, then why should we not conclude that this is a highly unusual occurrence? The answer to that question lies in the huge size of the bullet population and the relative insignificance of agent Peters’ own personal experience in such a huge population.²⁰⁸

In reaching this conclusion, Judge Guzmán recognized the long history of both FBI reliance on CBLA, and of federal courts admitting CBLA testimony at trial.²⁰⁹ In spite of this recognition, he followed the mandate of Rule 702 and excluded the unsupported testimony.²¹⁰

Within a year, the National Research Council had issued a report concluding that the CBLA discipline could not support statements about how unusual it was for bullets to be “analytically indistinguishable.”²¹¹ A year after that, the FBI discontinued its CBLA Laboratory.²¹² After more than thirty years of judges admitting CBLA evidence, its façade of scientific reliability crumbled rapidly.

cartridges in the box ‘likely originated from the same manufacturers’ source (melt) of lead.’”).

²⁰⁶ *Id.* at *6.

²⁰⁷ *Id.* at *5.

²⁰⁸ *Id.* A similar argument could be made for the assertion that all human mouths are unique. See *supra* note 34 and accompanying text.

²⁰⁹ *Mikos*, 2003 WL 22922197, at *6.

²¹⁰ *Id.*

²¹¹ CBLA REPORT, *supra* note 203, at 90–94, 106–08.

²¹² See CBLA Press Release, *supra* note 197 (announcing that the FBI “will no longer conduct the examination of bullet lead,” largely because “neither scientists nor bullet manufacturers are able to definitely attest to the significance of an association made between bullets in the course of a bullet lead examination”).

This analysis is not meant to suggest that a single, unpublished opinion from a district judge will lead to the imminent downfall of a forensic discipline. In fact, in the CBLA context, the National Research Council had begun meeting to prepare their report nearly a year before Judge Guzmán issued the *Mikos* ruling, and did so because CBLA had already “come under greater scrutiny.”²¹³ However, as the ruling to exclude CBLA evidence, *Mikos* was an important milestone in the path toward recognizing and addressing the forensic science problem in one context,²¹⁴ and demonstrates that judges can—and should—leverage Rule 702 to block bad evidence even in the face of a long history of admission.

CONCLUSION

Trial judges are uniquely well positioned to staunch the flow of unreliable forensic evidence into court, which will both prevent wrongful convictions and inspire scientific research to validate or improve forensic disciplines. In order to do so, however, trial judges must break with sometimes-lengthy histories of admission, engage in a technical analysis outside the wheelhouse of most lawyers, and perhaps even face political backlash against an unpopular decision. As difficult as this may seem, none of the obstacles facing trial judges are insurmountable, and none exempt trial judges from their obligation to vigilantly gatekeep expert evidence in their courtrooms.

Most importantly, if the impetus for change does not come from trial judges, the current political climate suggests that bad forensic evidence will continue to be admitted, and history tells us that wrongful convictions will follow. This Note has suggested a path forward for judges prepared to recognize and act on the forensic science problem. Now more than ever, trial judges must lead the way toward a better future for forensic evidence.

²¹³ CBLA REPORT, *supra* note 203, at ix.

²¹⁴ See Giannelli, *supra* note 200, at 310 (noting that *Mikos* was the first case to rule such evidence inadmissible).