

AMERICAN JOURNAL
of LAW and EQUALITY

POLICING FORENSIC EVIDENCE

Brandon L. Garrett*

INTRODUCTION

If there had been no video of Derek Chauvin killing George Floyd on May 25, 2020, the legal and public response might have been quite different. Before a series of videos captured by witnesses and security cameras surfaced, the Minneapolis Police Department issued a statement: “After Floyd got out of his car, he physically resisted officers. Officers were able to get the suspect into handcuffs and noted he appeared to be suffering medical distress.”¹ The Chief Medical Examiner for the county did determine that Floyd’s death was a homicide but also speculated that Floyd suffered from “significant conditions,” such as drug use and a heart condition, that may have increased his likelihood of death.² At Chauvin’s trial, the medical examiner testified, “In my opinion, the law enforcement subdual, restraint, and the neck compression was just more than Mr. Floyd could take by virtue of those heart conditions.”³ Attorneys for Floyd’s family commissioned an

*L. Neil Williams, Jr. Professor of Law, Duke University School of Law, and Director, Wilson Center for Science and Justice.

- 1 See *Investigative Update on Critical Incident*, INSIDE MINNEAPOLIS POLICE DEP’T (May 25, 2020), <https://web.archive.org/web/20200526183652/https://www.insidempd.com/2020/05/26/man-dies-after-medical-incident-during-police-interaction/>; see also Philip Bump, *How the First Statement from Minneapolis Police Made George Floyd’s Murder Seem Like George Floyd’s Fault*, WASH. POST (Apr. 20, 2021), <https://www.washingtonpost.com/politics/2021/04/20/how-first-statement-minneapolis-police-made-george-floyds-murder-seem-like-george-floyds-fault/>.
- 2 See HENNEPIN COUNTY MEDICAL EXAMINER’S OFFICE AUTOPSY REPORT (June 1, 2020), <https://www.hennepin.us/media/hennepinus/residents/public-safety/medical-examiner/floyd-autopsy-6-3-20.pdf> (citing, as the “case title,” “cardiopulmonary arrest complicating law enforcement subdual, restraint, and neck compression”); see also Bill Chappell & Laurel Wamsley, *Video: Medical Examiner Says Police Restraint ‘Just More Than Mr. Floyd Could Take,’* NPR (Apr. 9, 2021, 5:55 PM ET), <https://www.npr.org/sections/trial-over-killing-of-george-floyd/2021/04/09/985722945/live-video-medical-examiner-to-testify-about-george-floyds-death>.
- 3 Chappell & Wamsley, *supra* note 2.

© 2022 Brandon L. Garrett. Published under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International license (CC BY-NC-ND).
https://doi.org/10.1162/ajle_a_00042

independent autopsy that found that there were no such contributing factors.⁴ More specifically, they found that the cause of death was not a heart attack but rather “mechanical asphyxia”—meaning the flow of oxygen was cut off so Floyd, as he told the officers, simply could not breathe.⁵ As Martin Tobin testified at the criminal trial, given the force placed on Floyd’s neck, “[a] healthy person subjected to what Mr. Floyd was subjected to would have died.”⁶

It was not an aberration that the medical examiner’s determinations regarding cause of death were hotly disputed. There is a long, troubling history of prosecution, in media releases, labeling people killed by police as criminals or drug users or persons with preexisting medical conditions.⁷ Forensic evidence, in which medical examiners or coroners set out conclusions in death certificates, has played a troubling role in providing faulty support for such claims.⁸ All too often, questionable findings have been shared with the public and our courts, including by omitting entirely the role of police in cases of police killings.⁹ The coroner and medical examination field is rife with poor methods. A landmark 2009 National Academy of Sciences report (“NAS Report”) called death investigations in the United States “fragmented” as a system that is “deficient” in its practices and called for the replacement of local coroners’ offices with a medical examiner system.¹⁰

The Floyd case is just one high-profile example of a larger problem: police evidence collection, including when it is conducted with outside experts and crime labs, lacks adequate scientific safeguards and as a result is prone to bias and error.¹¹ A central function of policing is to gather evidence during criminal investigations. Law enforcement receives a great deal of deference when it carries out those investigative functions. While detailed constitutional criminal procedure rules regulate traditional policing functions, such as searches and seizures, there is very little regulation of forensic evidence of any kind, much less regulation informed by community priorities, research, or policy.¹² To be sure, a growing number of police

4 See Bill Chappell & Vanessa Romo, *Chauvin Trial: Expert Says George Floyd Died from a Lack of Oxygen, Not Fentanyl*, NPR (Apr. 8, 2021, 6:18 PM ET), <https://www.npr.org/sections/trial-over-killing-of-george-floyd/2021/04/08/985347984/chauvin-trial-medical-expert-says-george-floyd-died-from-a-lack-of-oxygen>.

5 *Id.*

6 *Id.*

7 See Sam Levin, *Killed by Police, Then Vilified: How America’s Prosecutors Blame Victims*, GUARDIAN, Mar. 21, 2019; see also Samantha Michaels, *Why Coroners Often Blame Police Killings on a Made-Up Medical Condition*, MOTHER JONES, Oct. 14, 2020 (discussing attributions of deaths to drug overdose or excited delirium).

8 See Tim Arango & Shaila Dewan, *More Than Half of Police Killings Are Mislabeled, New Study Says*, N.Y. TIMES (Sept. 30, 2021), <https://www.nytimes.com/2021/09/30/us/police-killings-undercounted-study.html>.

9 *Id.*

10 See COMM. ON IDENTIFYING THE NEEDS OF THE FORENSIC SCIS. CMTY. & NAT’L RSCH. COUNCIL, STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD 29 (2009) [hereinafter NAS REPORT].

11 For a much more comprehensive account, see BRANDON L. GARRETT, *AUTOPSY OF A CRIME LAB: EXPOSING THE FLAWS IN FORENSICS* (2020).

12 Paul C. Giannelli, *Wrongful Convictions and Forensic Science: The Need to Regulate Crime Labs*, 86 N.C. L. REV. 163 (2007).

policies and guidelines reflect the recent effort to emphasize accurate and impartial evidence collection. The International Association of Chiefs of Police has taken an active role in promoting the consideration of ways to improve the “accuracy and thoroughness” of police investigations.¹³ Far more is needed to ensure such accuracy and thoroughness.

A second example of forensics gone wrong illustrates the problem, not in the medical examiner setting but in the crime laboratory setting. The Department of Forensic Science (DFS) in Washington, D.C., was supposed to be a model lab that would change lab work on crime scene evidence for the better. The D.C. government passed the Department of Forensic Sciences Act of 2011 to make the lab more independent of police and require it to be accredited.¹⁴ It paid tens of millions of dollars for a large new facility.¹⁵ Unfortunately, while accredited crime labs may have more in the way of quality control than a typical policing agency, they do not have the quality controls present in a scientific or clinical lab.

In a murder trial, Judge Todd Edelman ruled in August 2019 that after hearing from several experts at an extensive hearing, he was concerned that published studies had not been able to show that DFS examiners could reliably make a “source attribution.”¹⁶ The judge held that the most an expert could say for the government was that “the recovered firearm [could not] be excluded as the source of the cartridge casing found on the scene of the alleged shooting.”¹⁷ A new case involving firearms evidence analyzed in 2017 and charges against two men for two killings was in front of Judge Edelman.¹⁸ The firearms group at the lab had reported that the same weapon fired the cartridges in the two shootings.¹⁹ Perhaps because Judge Edelman had been skeptical in the past, the prosecutors wanted to make sure; they took an unusual step and asked an independent examiner to take a look at the evidence.²⁰ That examiner was sure that it was clear that two different

13 INTERNATIONAL ASS'N OF CHIEFS OF POLICE, NATIONAL SUMMIT ON WRONGFUL CONVICTIONS: BUILDING A SYSTEMIC APPROACH TO PREVENT WRONGFUL CONVICTIONS 10 (2013).

14 See D.C. CODE § 5-1501.02 (West 2022) (“There is established as a subordinate agency in the executive branch of the government of the District of Columbia, the Department of Forensic Sciences.”).

15 Editorial Board, *D.C.’s New Crime Lab Goes Under the Microscope*, WASH. POST (Mar. 11, 2015), https://www.washingtonpost.com/opinions/dcs-new-crime-lab-goes-under-the-microscope/2015/03/11/fe48dfd2-c442-11e4-ad5c-3b8ce89f1b89_story.html.

16 Memorandum Opinion at *1, *Tibbs v. United States*, No. 2016 CF1 19431 (D.C. Super. Ct. Sept. 5, 2019), <https://context-cdn.washingtonpost.com/notes/prod/default/documents/cc85da89-f6a1-4172-bf1c-b6b759669687/note/2faab6e6-85da-4abe-a669-b9f48db2498e.pdf> (“According to the government’s proffer, this analysis permitted the examiner to identify the recovered firearm as the source of the cartridge casing collected from the scene.”).

17 *Id.* (“[T]he government’s expert witness must limit his testimony to a conclusion that, based on his examination of the evidence and the consistency of the class characteristics and microscopic toolmarks, the firearm cannot be excluded as the source of the casing.”).

18 Jack Moore, *DC Judge Orders Forensic Lab to Turn Over Some Documents Sought by Prosecutors*, WTOP NEWS (Dec. 10, 2020, 2:34 PM), <https://wtop.com/dc/2020/11/dc-judge-orders-forensic-lab-to-turn-over-some-documents-sought-by-prosecutors/>.

19 *Id.*

20 *Id.*

firearms were involved.²¹ The lab defiantly insisted the outside examiner was wrong and that in fact the examiner’s report was an “administrative error” and the evidence was now deemed to be “inconclusive.”²²

The lab went into damage control mode. The lab director called it a “darn good lab.”²³ Management told the American National Standards Institute National Accreditation Board, which accredited the lab, that it had made an “inconclusive” finding after initiating a review of the incident. In April 2020, the lab’s accreditation was suspended.²⁴ The lab entirely stopped doing work as a result. Prosecutors opened a new probe into the firearms unit and disagreed with the lab in eleven cases. A scathing audit report found that lab managers sought to conceal the problem in the case before Judge Edelman and pressured others to report that the finding was “inconclusive.”²⁵ One examiner recalled how he tried to “push back as long as he could” and wished he had resigned in protest.²⁶ The auditors concluded that this illustrated “very serious” problems with management that indicated lack of adherence to “core principles of integrity” and ethics. The lab director resigned. A letter alleged “potential concealment” of errors in the fingerprint unit as well.²⁷ In September 2021, the lab disbanded the firearms unit. A criminal investigation of the lab is pending, and the lab remains closed.²⁸

As these high-profile examples both show, forensic evidence has, to varying degrees, left the police station for the laboratory in a manner that represents a highly incomplete and much-needed transition to sound science. In recent years, the scientific community has called for forensic work to be independent of police so it can adhere to sound scientific standards. The central recommendation of the NAS Report, which brought together leading scientists to consider the path forward for forensic science, centered on independence and scientific methods.²⁹ At best, some crime labs have obtained financial but not

21 *Id.*

22 *Id.*

23 Jack Moore & Megan Cloherty, ‘You Can Trust This Laboratory’: DC Crime Lab Director Responds to Scrutiny of Firearms Unit, WTOP NEWS (Dec. 2, 2020, 4:24 AM), <https://wtop.com/dc/2020/12/you-can-trust-this-laboratory-dc-crime-lab-director-responds-to-scrutiny-of-firearms-unit/>.

24 Keith L. Alexander, *National Forensic Science Board Suspends D.C. Crime Lab’s Accreditation, Halting Analysis of Evidence, City Says*, WASH. POST (Apr. 3, 2021), https://www.washingtonpost.com/local/public-safety/dc-lab-forensic-evidence-accreditation/2021/04/03/723c4832-94aa-11eb-a74e-1f4cf89fd948_story.html.

25 Praeipce, *U.S. v. McLeod*, No. 2017 CF1 9869 (D.C. Super. Ct. Mar. 22, 2021), https://wtop.com/wp-content/uploads/2021/03/rondell_mcleod_dfs_court_filing_march_22.pdf.

26 *Id.*

27 Paul Wagner, *DC Crime Lab Under Investigation After Allegations of Wrongdoing*, NBC NEWS (Apr. 8, 2021, 8:40 PM), <https://www.nbcwashington.com/news/local/dc-crime-lab-under-investigation-after-allegations-of-wrongdoing/2634489/>.

28 Jack Moore, *DC Abruptly Disbands Crime Lab’s Firearms Unit*, WTOP NEWS (Sept. 16, 2021, 4:00 PM), <https://wtop.com/dc/2021/09/dc-abruptly-disbands-crime-labs-firearms-unit/>.

29 See NAS REPORT, *supra* note 10.

structural independence from law enforcement with the accountability and accuracy that should accompany an embrace of scientific methods.

Many are asking new questions about how to define public safety.³⁰ Even in the area of forensic evidence, where there is a greater expectation that science will be used in service of public safety, questions about the police's role in collecting and testing evidence have remained unasked. Far too often, untrained police, rather than trained scientists, collect and potentially alter or contaminate evidence at crime scenes.³¹ Further, we allow law enforcement oversight of most crime labs, or we allow law enforcement and prosecutors to direct and influence their work, practices that would never be tolerated in a scientific experiment, much less a scientific lab.³² As police deploy new technologies, these questions will hopefully continue to be asked, but little scrutiny has resulted when old technologies, like pathology examinations or fingerprint comparisons, have been used. One reason these practices remain under-examined is that neither courts nor the public can readily access what procedures crime labs employ or what results they obtain.

However, the movement to bring scientific standards to forensics, which is increasingly embraced by the forensic community and legal actors, has the potential to address these long-standing needs. As I discuss in this essay, that movement also has promise for policing more generally. We should be using quality controls to regulate crime labs, and policing more broadly, to evaluate the fairness and public safety of outcomes. The focus here is on three specific core principles that can improve the role of forensic evidence in policing: (1) independence, (2) accuracy, and (3) oversight.

I. INDEPENDENCE AND FORENSIC EVIDENCE

The NAS Report warned that a “lack of independence” of crime labs from law enforcement can damage the objectivity of forensic science and harm criminal justice.³³

30 Daniel S. Nagin, *Cost-Benefit Analysis of Crime Prevention Policies*, 14 *CRIMINOLOGY & PUB. POL'Y* 583, 585 (2015); Barry Friedman & Elizabeth Jansky, *Policing's Information Problem* (2019), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3452242; Barry Friedman, *We Spend \$100 Billion on Policing. We Have No Idea What Works.*, *WASH. POST* (Mar. 10, 2017), <https://www.washingtonpost.com/posteverything/wp/2017/03/10/we-spend-100-billion-on-policing-we-have-no-idea-what-works/>.

31 For an in-depth analysis of this problem, see Jennifer E. Laurin, *Remapping the Path Forward: Toward A Systemic View of Forensic Science Reform and Oversight*, 91 *TEX. L. REV.* 1051, 1081 (2013); see also FRANK HORVATH ET AL., A NATIONAL SURVEY OF POLICE POLICIES AND PRACTICES REGARDING THE CRIMINAL INVESTIGATION PROCESS: TWENTY-FIVE YEARS AFTER RAND 76 (2001), <https://www.ncjrs.gov/pdffiles1/nij/grants/202902.pdf> (“[I]n most agencies evidence-related duties are not assigned predominantly to any one type of individual or position. Rather, they are more likely to be shared among patrol officers . . . investigators . . . and evidence technicians . . .”).

32 For the authoritative account, see SANDRA GUERRA THOMPSON, *COPS IN LAB COATS: CURBING WRONGFUL CONVICTIONS THROUGH INDEPENDENT FORENSIC LABORATORIES* 181–83 (2015).

33 See NAS REPORT, *supra* note 10.

Traditionally, forensic evidence has largely been a law enforcement function, not a scientific function, in the United States. Early crime labs were small and operated by police departments, following the early example of the Federal Bureau of Investigation (FBI) lab.³⁴ By the 1960s, every state had crime labs run by law enforcement.³⁵ Today, there are over four hundred publicly funded crime labs with over 14,300 lab employees.³⁶ Very few labs in the United States are independent from law enforcement.

Nor is the work done by professionals necessarily independent of law enforcement. This problem begins at the crime scene. Such work may be conducted by a variety of professionals and roles, including officers working for police departments but also crime-lab analysts, technicians, and others.³⁷ Evidence collection may be conducted by law enforcement. Law enforcement priorities and resources may affect the initial decision whether to test evidence. Recent inquiries and audits have resulted when labs have failed to test forensic evidence, particularly regarding untested rape kits. One study found that forty percent of unanalyzed rape and homicide cases were estimated to have testable DNA.³⁸ In Los Angeles County, following a 2009 Human Rights Watch report and public protests, police and sheriffs began to work to address years of backlogs.³⁹ In response to the serious problem of non-testing, and with federal grant support,⁴⁰ twenty-six states have enacted statutes requiring inventories of untested sexual assault evidence.⁴¹

Police begin the process of examining forensic evidence when they collect evidence at the crime scene. Police often do not have policies regarding sound and scientific collection of crime scene evidence, which then affects the work done by a lab.⁴² There has been a recent focus on developing policies and procedures for crime scene evidence collection,

34 Paul Giannelli, *Forensic Science: Why No Research?*, 38 FORDHAM URBAN L.J. 503 (2010).

35 See John I. Thornton, *Criminalistics: Past, Present and Future*, 11 LEX ET SCIENTIA 23 (1975).

36 MATTHEW R. DUROSE ET AL., BUREAU OF JUST. STAT., U.S. DEP'T OF JUST., CENSUS OF PUBLICLY FUNDED FORENSIC CRIME LABORATORIES, 2009, at 1 (2012), <https://bjs.ojp.gov/content/pub/pdf/cpffcl09.pdf>.

37 See JOSEPH PETERSON & IRA SOMMERS, THE ROLE AND IMPACT OF FORENSIC EVIDENCE IN THE CRIMINAL JUSTICE PROCESS 22 (2010).

38 KEVIN J. STROM ET AL., 2007 SURVEY OF LAW ENFORCEMENT EVIDENCE PROCESSING: FINAL REPORT 4–5 (2009).

39 HUM. RTS. WATCH, TESTING JUSTICE: THE RAPE KIT BACKLOG IN LOS ANGELES CITY AND COUNTY (2009), <https://www.hrw.org/sites/default/files/reports/rapekit0309web.pdf>.

40 The SAFER Act of 2013, Pub. L. No. 113-4, § 1002, 127 Stat. 54, 127–131 (authorizing use of grants under the Debbie Smith DNA Backlog Grant Program (34 U.S.C. § 40701) to conduct audits of sexual assault evidence and requiring the Attorney General to publish information from audits online); see also *Sexual Assault Kit Initiative (SAKI)*, BUREAU OF JUST. ASSISTANCE, DEP'T OF JUST. PROGRAM, https://www.bja.gov/ProgramDetails.aspx?Program_ID=117 (last visited Dec. 2018).

41 U.S. GOV'T ACCOUNTABILITY OFF., GAO-19-216, DNA EVIDENCE: DOJ SHOULD IMPROVE PERFORMANCE MEASUREMENT AND PROPERLY DESIGN CONTROLS FOR NATIONWIDE GRANT PROGRAM 16 (2019), <https://www.gao.gov/assets/700/697768.pdf>.

42 HORVATH ET AL., *supra* note 31; Laurin, *supra* note 31, at 1081.

including to promote the use of forensic professionals at the crime scene.⁴³ Not only do police often lack the resources or training to adequately collect evidence at the crime scene, but they often fail to collect evidence that could be highly probative.⁴⁴ Police may also preemptively test evidence themselves using inaccurate field methods. For example, commercial drug-testing kits are often not carefully validated; the few studies conducted on this issue have shown that these kits can have high error rates. The field tests are supposed to be followed up with a more rigorous lab test, but in the meantime, a person may be arrested and face great pressure to plead guilty. In Harris County, Texas, an audit by the District Attorney's Office Conviction Integrity Unit uncovered hundreds of such cases. The Texas Forensic Science Commission later found these tests too unreliable for use and the convictions were subsequently reversed.⁴⁵ In 2017, Houston police ceased use of field drug tests, but they remain widely used in other jurisdictions.

In other areas, bias in police evidence collection can be of constitutional concern. For example, the U.S. Supreme Court has found highly suggestive and unreliable practices in the area of eyewitness identifications that may violate the Due Process Clause. The Court's ruling in *Manson v. Brathwaite* set out a due process test using a series of so-called reliability factors to assess eyewitness identifications in which police used suggestive procedures.⁴⁶

No such inquiry exists for forensic evidence. A large body of research has now documented how cognitive biases can affect forensic work. Psychological research has shown how where forensic techniques involve some degree of judgment and interpretation, experts are vulnerable to cognitive bias.⁴⁷ In the laboratory, testing is rarely blind, and a range of communications with law enforcement can bias the test results. Traditionally, beginning with the submission of evidence, the forensic analysts communicate with law enforcement about evidence in the case and hear task-irrelevant and biasing information.

43 See, e.g., NAT'L FORENSIC SCI. TECH. CTR. (NFSTC), CRIME SCENE INVESTIGATION: A GUIDE FOR LAW ENFORCEMENT (2014), <https://shop.nfstc.org/crime-scene-investigation-guide/>; SUE BALLOU ET AL., THE BIOLOGICAL EVIDENCE PRESERVATION HANDBOOK: BEST PRACTICES FOR EVIDENCE HANDLERS (2013), <https://www.nist.gov/system/files/documents/forensics/NIST-IR-7928.pdf>.

44 See PETERSON & SOMMERS, *supra* note 37, at 22.

45 Ryan Gabrielson, *Texas Panel on Wrongful Convictions Calls for Ending Use of Unverified Drug Field Tests*, PROPUBLICA (Jan. 18, 2017, 12:48 PM EST), <https://www.propublica.org/article/texas-panel-wrongful-convictions-calls-end-use-unverified-drug-field-tests>.

46 *Manson v. Brathwaite*, 432 U.S. 98, 114 (1977).

47 EXEC. OFF. OF THE PRESIDENT, PRESIDENT'S COUNCIL OF ADVISORS ON SCI. & TECH. (PCAST), FORENSIC SCIENCE IN CRIMINAL COURTS: ENSURING SCIENTIFIC VALIDITY OF FEATURE-COMPARISON METHODS 8 (2016), https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf [hereinafter PCAST REP.] ("The findings of forensic science experts are vulnerable to cognitive and contextual bias.").

In one of the most famous studies in all of forensics, five highly experienced fingerprint examiners reviewed a set of prints in the course of their ordinary work.⁴⁸ The examiners did not know that Professor Itiel Dror and colleagues had given these examiners prints that each of the five had looked at years before in a real-life case and had found then to be from the same source.⁴⁹ Other experts had separately looked at those prints and agreed that they were from the same source. Dror and colleagues informed the five examiners that the prints were from the infamous Madrid bombing case, in which three senior FBI examiners mistakenly implicated an innocent Portland, Oregon, lawyer.⁵⁰ After having been told that these prints were from that case, in which such a high-profile error had been made, four of the five reached a different conclusion: one persisted in reporting that the prints came from the same source, three now concluded these were “definite non-match” prints, and the fourth called the evidence “inconclusive.”⁵¹

Addressing cognitive bias is only now becoming a central part of quality control at crime laboratories. Adoption of procedural protections can reduce such bias. For example, a linear sequential approach has been developed for comparative pattern disciplines. The 2016 President’s Council of Advisors on Science and Technology (PCAST) Report recommends that approach for fingerprint comparisons: an examiner “should be required to complete and document their analysis of a latent fingerprint before looking at any known fingerprint” to prevent circular reasoning where looking back and forth between prints encourages an expert to disregard differences.⁵² A second protection is to ensure that potentially biasing task-irrelevant information is not passed on to lab analysts.⁵³ The PCAST Report also recommended that a fingerprint examiner “should separately document any additional data used during their comparison and evaluation.”⁵⁴ Information is task-irrelevant if it is not necessary for drawing conclusions about the propositions in question, using the method in question. Separation of roles in a laboratory can help keep task-irrelevant information from biasing analysts.⁵⁵ Such procedures are much needed to define

48 I. E. Dror et al., *Contextual Information Renders Experts Vulnerable to Making Erroneous Identifications*, 156 *FORENSIC SCI. INT’L.* 74, 77 (2006). (“Our study shows that it is possible to alter identification decisions on the same fingerprint, solely by presenting it in a different context.”).

49 *Id.*

50 *Id.*

51 *Id.*

52 See PCAST REP., *supra* note 47, at 10; see also Dan E. Krane et al., *Sequential Unmasking: A Means of Minimizing Observer Effects in Forensic DNA Interpretation*, 53 *J. FORENSIC SCI.* 1006–07 (2008); Itiel E. Dror et al., *Context Management Toolbox: A Linear Sequential Unmasking (LSU) Approach for Minimizing Cognitive Bias in Forensic Decision Making*, 60 *J. FORENSIC SCI.* 1111 (2015).

53 See NAT’L COMM’N ON FORENSIC SCI., *ENSURING THAT FORENSIC ANALYSIS IS BASED UPON TASK-RELEVANT INFORMATION 2* (2015), <https://www.justice.gov/archives/ncfs/page/file/641676/download>.

54 See PCAST REP., *supra* note 47, at 10.

55 See Itiel E. Dror, *A Hierarchy of Expert Performance*, 5 *J. APPLIED RSCH. IN MEMORY AND COGNITION* 121–27 (2016).

the relationships between law enforcement and crime labs, and they should be a hallmark of functionally independent and sound science.⁵⁶ Those reforms are only slowly being implemented, however.

II. ACCURACY

A wide range of scientific and legal actors, including scientists, forensic practitioners, lawmakers, lawyers, and judges, have called for greater reliability in forensics. The National Academy of Sciences Report in 2009 emphasized that “some forensic science disciplines are supported by little rigorous systematic research to validate the discipline’s basic premises and techniques.” Slowly, as such research has begun to be conducted, we have learned just how often forensic methods go wrong.⁵⁷

The legal response has been relatively piecemeal. Many states now follow standards based on the U.S. Supreme Court’s ruling in *Daubert v. Merrell Dow Pharmaceuticals*⁵⁸ and its progeny, which task judges with greater gatekeeping responsibilities to assess the reliability of scientific evidence. Those rules are widely understood to be ineffectively used in criminal cases. Leading reports, such as the influential 2009 National Academy of Sciences Report, have noted that lawyers and judges have not taken an active or effective role in reviewing the reliability of forensic evidence in criminal cases.⁵⁹ Constitutional criminal procedure bars the failure to disclose exculpatory evidence and the outright fabrication of evidence by law enforcement. The U.S. Supreme Court has long held that fabricating evidence or knowingly using perjured testimony violates the Due Process Clause.⁶⁰ In addition, officers and prosecutors must supply “any favorable evidence known to others acting on the government’s behalf in the case, including the police.”⁶¹ No constitutional rules regulate the reliability of forensic evidence, apart from those elemental disclosure and non-fabrication rules.

For decades, forensic analysts of different types testified they were one hundred percent certain of their conclusions.⁶² As federal judge Harry T. Edwards put it, “The courts had been misled for a long time because we had been told, my colleagues and I, by some

56 For an example of a simple checklist that can assist with such work, see Adele Quigley-McBride et al., *A Practical Tool for Information Management in Forensic Decisions: Using Linear Sequential Unmasking-Expanded (LSU-E) in Casework*, 4 FORENSIC SCI. INT’L: SYNERGY 100216 (2022).

57 See PCAST REP., *supra* note 47, at 56 (discussing black box studies conducted for certain disciplines).

58 *Daubert v. Merrell Dow Pharms.*, 509 U.S. 579 (1993).

59 See NAS Report, *supra* note 10; Peter J. Neufeld, *The (Near) Irrelevance of Daubert to Criminal Justice and Some Suggestions for Reform*, 95 AM. J. PUB. HEALTH S107, S110 (2005).

60 *Napue v. Illinois*, 360 U.S. 264 (1959); *Mooney v. Holohan*, 294 U.S. 103 (1935).

61 *Kyles v. Whitley*, 514 U.S. 419, 437 (1995).

62 See Alex Biedermann et al., *After Uniqueness: The Evolution of Forensic-Science Opinions*, JUDICATURE, Spring 2018, <https://judicature.duke.edu/articles/after-uniqueness-the-evolution-of-forensic-science-opinions/>.

experts from the FBI that fingerprint comparisons involved essentially a zero error rate, without our ever understanding that's completely inaccurate.”⁶³ Yet, no one had carefully tested the basic assumptions that experts have relied upon for decades. More recently, the NAS and PCAST Reports highlighted the need for scientific testing of the reliability of core forensic disciplines. The PCAST Report detailed how for a more objective method, like a drug test, you can evaluate each step in the process by seeing whether it produces accurate results.⁶⁴ However, for subjective techniques like pattern evidence comparisons, there are no clearly defined steps. The person is the process: an examiner's mind is a “black box” that reaches judgments based on training and experience.⁶⁵ To test such a “black box” examiner, one can give tests with samples of realistic difficulty and with the correct answer known in advance.⁶⁶ In forensics, that type of proficiency testing is not routine in labs, as part of quality control or performance testing of examiners, except using very elementary tests that are not designed to be rigorous.⁶⁷ Entire fields have lacked such testing to establish reliability. And, as the PCAST Report underscored, “without appropriate estimates of accuracy, an examiner's statement that two samples are similar—or even indistinguishable—is scientifically meaningless: it has no probative value, and considerable potential for prejudicial impact.”⁶⁸ For example, the PCAST Report described that, while several early studies had been conducted and documented high error rates, just two properly designed studies of the accuracy of latent fingerprint analysis had been conducted to date.⁶⁹ That alone is deeply disturbing, given the prominence of fingerprint evidence in criminal investigations and prosecutions over the past hundred years. It was generous for the report to say that just two studies would be enough to permit a technique to be used in criminal cases. Indeed, the people participating in the studies knew that they were being tested and that it was important for the field. They were likely very cautious in their work, and they could drop out of the study, including if they found a question challenging and they feared making an error. In the larger of the two studies—the only one published—a huge group, eighty-five percent of the examiners, made at least one false negative error.⁷⁰

63 See Lowell Bergman, *The Real CSI*, FRONTLINE, <https://www.pbs.org/wgbh/frontline/film/real-csi/transcript/> (last visited June 23, 2022).

64 See PCAST REP., *supra* note 47, at 56.

65 *Id.* at 5–6 (stating that for subjective feature comparison methods, because the individual steps are not objectively specified, the method must be evaluated as if it were a “black box” in the examiner's head).

66 *Id.* at 6.

67 See Brandon L. Garrett & Gregory Mitchell, *The Proficiency of Experts*, 166 U. PA. L. REV. 901 (2018).

68 PCAST REP., *supra* note 47, at 6.

69 See *id.* at 98.

70 See Bradford T. Ulery et al., *Accuracy and Reliability of Forensic Latent Fingerprint Decisions*, 108 PROC. NAT'L ACAD. SCIS. 7733 (2011) (“Eighty-five percent of examiners made at least one false negative error for an overall false negative rate of 7.5%.”).

These findings provided another overdue wake-up call. It would shock jurors to hear of such error rates, and it would affect their convictions.⁷¹

The 2009 National Academy of Sciences Report similarly concluded that there needs to be more research “to confirm the fundamental basis for the science of bite mark comparison.”⁷² It said that the uniqueness of human dentition has “not been scientifically established.”⁷³ The scientists who wrote the PCAST Report concluded that since no valid studies of error rates had been done, the analysis lacked validity and should not be used.⁷⁴ What we do know about reliability is disturbing, as the PCAST Report detailed.⁷⁵ None of these troubling findings blunted the force of standard testimony on bite marks delivered in court, nor did forensic dentists make a habit of describing these studies in reports or testimony; indeed, courts continue to allow the evidence in criminal trials.⁷⁶

Of particular importance to police investigations, in part because firearms violence is a major problem in the United States, is that firearms comparisons are in great demand. To complete such analysis, examiners seek to link crime scene evidence, such as spent shell casings or bullets, with a firearm.⁷⁷ The assumption is that manufacturing processes leave markings on the barrel, breech face, and firing pin. When the firearm discharges, those components contact the ammunition and leave marks on it.⁷⁸ Experts have assumed that different firearms should leave different toolmarks on the ammunition such that the spent ammunition can be definitively linked to a single firearm.⁷⁹ By the late 1990s, experts premised testimony on a “theory of identification” set out by a professional association, the Association of Firearms and Tool Mark Examiners (AFTE), which instructs practitioners to use the phrase “source identification” to explain what they mean when they identify “sufficient agreement” when examining firearms.⁸⁰ The AFTE’s theory is circular: an identification occurs when the expert finds sufficient evidence defined as sufficient to find an identification.⁸¹ In recent years, scientists have called into question the validity and

71 See Brandon L. Garrett et al., *Error Rates, Likelihood Ratios, and Jury Evaluation of Forensic Evidence*, 65 J. FORENSIC SCIS. 1199 (2020).

72 See NAS Report, *supra* note 10.

73 *Id.*

74 See PCAST REP., *supra* note 47, at 87 (“PCAST finds that bitemark analysis does not meet the scientific standards for foundational validity, and is far from meeting such standards. To the contrary, available scientific evidence strongly suggests that examiners cannot consistently agree on whether an injury is a human bitemark and cannot identify the source of bitemark with reasonable accuracy.”).

75 See PCAST REP., *supra* note 47, at 87 (“Among those studies that have been undertaken, the observed false positive rates were so high that the method is clearly scientifically unreliable at present.”).

76 See GARRETT, *AUTOPSY OF A CRIME LAB*, *supra* note 11, at 127–29, 197.

77 See PCAST REP., *supra* note 47, at 104.

78 *Id.*

79 *Id.*

80 *Id.*

81 *Id.*

reliability of such testimony. In a 2008 report on ballistics imaging, the National Academy of Sciences (NAS) concluded that definitive associations like “source identification” were not supported.⁸² In its 2009 report, the NAS followed up and stated that categorical conclusions regarding firearms or toolmarks were not supported by research and that, instead, more cautious claims should be made.⁸³ The report stated that the “scientific knowledge base for tool mark and firearms analysis is fairly limited.”⁸⁴ The AFTE theory of identification “is inadequate and does not explain how an expert can reach a given level of confidence in a conclusion.”⁸⁵

By 2016, only a single black box study on firearms comparisons had been done, and it showed an error rate as high as 1 in 46.⁸⁶ This single study is unpublished, and it raises many of the same concerns as other studies in which participants were aware that their assignment was a test and in which some could and did drop out after reviewing the materials.⁸⁷ The rate of inconclusive errors was over thirty-three percent.⁸⁸ To this day, firearms examiners use terms like “source identification” in court—and avoid reporting what is known about error rates. Some judges have begun to step in and require more cautious wording for firearms conclusions.⁸⁹ However, crime lab regulation is needed, as described next, just as other scientific laboratories are regulated.

III. OVERSIGHT OF CRIME LABS

There are some ways in which crime labs can be a model for policing agencies. Forensic evidence should bring with it better documentation, scientific methods, and quality controls, which are broadly needed in the criminal system. Forensic examiners must follow lab protocols to document their work; police detectives and patrol officers should be held to higher standards regarding documentation of their investigative work. Forensic labs may be accredited and are required to adopt certain quality control standards. In addition,

82 See NAT'L RSCH. COUNCIL, BALLISTIC IMAGING 3–4 (2008).

83 See NAS Report, *supra* note 10, at 154.

84 *Id.*

85 *Id.*

86 See PCAST REP., *supra* note 47, at 11 (“The scientific criteria for foundational validity require that there be more than one such study, to demonstrate reproducibility, and that studies should ideally be published in the peer-reviewed scientific literature. Accordingly, the current evidence still falls short of the scientific criteria for foundational validity.”).

87 See DAVID P. BALDWIN ET AL., DEF. BIOMETRICS & FORENSICS OFF., ASSISTANT SEC’Y OF DEF., A STUDY OF FALSE-POSITIVE AND FALSE-NEGATIVE ERROR RATES IN CARTRIDGE CASE COMPARISONS (Technical Report #IS-5207) 9 (2014), <https://afte.org/uploads/documents/swggun-false-positive-false-negative-usdoe.pdf>.

88 See PCAST REP., *supra* note 47, at 110.

89 Regarding the effectiveness of such measures, see Brandon L. Garrett et al., *Mock Jurors’ Evaluation of Firearm Examiner Testimony*, 44 LAW & HUM. BEHAV. 412 (2020).

forensic examiners are subjected to periodic proficiency testing, which, in theory, examines the quality of their work. However, those tests are understood to be elementary and are often not taken individually or under realistic conditions.⁹⁰ Crime labs offer a model for more accurate policing, even though they themselves traditionally have fallen short of the independence, oversight, and quality controls in clinical and scientific laboratories outside the criminal system. Far more rigorous regulation of crime labs and forensic evidence is needed, and such a model offers lessons for how to improve policing more generally.

The National Academy of Sciences summarized the state of affairs facing forensics in the United States in 2009: “Forensic science facilities exhibit wide variability in capacity, oversight, staffing, certification, and accreditation across federal and state jurisdictions.”⁹¹ “Rigorous mandatory certification and accreditation programs, adherence to robust performance standards, and effective oversight”⁹² are lacking. Moreover, what policies labs do have may be nonpublic.⁹³

In contrast, public regulations for medical laboratories have developed over time to support public health goals. After World War II, medical laboratories that conducted an experiment to assess clinical labs in Pennsylvania found serious errors, such as failure to correctly identify diseases. In 1967, Congress passed the Clinical Laboratory Improvement Act (CLIA) to ensure that medical labs conduct accurate tests.⁹⁴ A second wave of reform began in the mid-1980s, after reporters at the *Wall Street Journal* wrote about misdiagnosed cancer and lax standards at labs in a Pulitzer Prize-winning series.⁹⁵ In 1988, Congress passed amendments to the CLIA extending the regulations to practically all clinical laboratories, public or not. The law required that proficiency testing reflect “to the extent practicable . . . normal working conditions” to make tests realistic.⁹⁶ The law also permitted “announced and unannounced on-site proficiency testing of such individuals.”⁹⁷ After all, the lawmakers concluded, “regular proficiency testing [is] vital evidence of a laboratory’s competence.”⁹⁸

90 See PCAST REP., *supra* note 47, at 57 (stating that Christopher Czyryca, the president of Collaborative Testing Services, Inc., the leading proficiency testing firm in the U.S., has publicly said that “[e]asy tests are favored by the community”).

91 NAS Report, *supra* note 10, at 14.

92 *Id.* at 6.

93 Sandra Guerra Thompson & Nicole Cásarez, *Three Transformative Ideals to Build a Better Crime Lab*, 34 GA. ST. U. L. REV. 1007 (2018).

94 See 42 U.S.C. § 263a (2012).

95 See Walt Bogdanich, *Lax Laboratories: The Pap Test Misses Much Cervical Cancer Through Labs’ Errors*, WALL ST. J., Nov. 2, 1987, at A1.

96 42 U.S.C. § 263a(f)(4)(B)(iv).

97 See 42 U.S.C. § 263a(f)(4)(B)(iv) (2012).

98 See H.R. Rep. No. 100-899, at 11 (1988), reprinted in U.S.S.C.A.N. at 3831; S. Rep. No. 100-561, at 3-4 (1988).

The swift responses by federal lawmakers to clinical lab failures was completely different from the tepid responses to crime lab failures. We need similarly serious federal legislation and quality controls imposed on crime labs. At the state level, there are a few state forensic science commissions, but only the Texas Forensic Science Commission has done audits and investigations in even a modest way.⁹⁹ Maryland began licensing crime labs in 2007 after a discredited state police ballistics and firearms expert was found to have falsified his academic credentials, resulting in a review of over 4,000 cases.¹⁰⁰ In response, Maryland adopted a model based on the CLIA, extending clinical regulations to crime laboratories.¹⁰¹ The statute and accompanying regulations provide a model for detailed regulation of crime laboratories; however, the effectiveness of the regulatory body that enforces the statute and accompanying regulations is unclear, and problems have persisted, including in firearms examinations, which had originally prompted the regulations.¹⁰²

As a result of the nonexistence of federal quality controls, many crime laboratories have long “lacked quality control measures that would have detected” uses of scientifically “questionable evidence.”¹⁰³ No federal or national accreditation system exists in the United States. Some states have required that their labs be accredited; otherwise, accreditation is voluntary.¹⁰⁴ The National Commission on Forensic Science strongly recommends that all forensic science service providers (FSSPs) become accredited to promote compliance with industry best practices, promote standardization, and improve quality of services.¹⁰⁵ The American Bar Association, similarly, has made this recommendation: “Crime laboratories and medical examiner officers should be accredited, examiners should be certified, and procedures should be standardized and published to ensure the validity, reliability, and timely analysis of forensic evidence.”¹⁰⁶ Accreditation has become far more common among

99 For an overview, see GARRETT, *AUTOPSY OF A CRIME LAB*, *supra* note 11, at 201–03.

100 Steve Lash, *Maryland High Court Lifts Attorneys’ Diligence Burden in Fraudulent Ballistics Expert Appeals*, MD. DAILY REC. (June 9, 2021), <https://thedailyrecord.com/2021/06/09/md-high-court-lifts-attorneys-diligence-burden-in-fraudulent-ballistics-expert-appeals/>; Jennifer McMenamin, *Police Expert Lied About Credentials*, BALT. SUN (Mar. 9, 2007, 3:00 AM), <https://www.baltimoresun.com/maryland/bal-te.md.forensics09mar09-story.html>.

101 MD. CODE ANN., HEALTH–GEN. § 17-2A (West 2022); for accompanying regulations, see MD. CODE REGS. 10.51.00-07.

102 See, e.g., Justin Fenton, ‘*Serious Questions’ Raised by Reports on Problems Inside Baltimore Police Crime Lab*, BALT. SUN (Aug. 16, 2021, 2:18 PM), <https://www.baltimoresun.com/news/crime/bs-md-ci-cr-crime-lab-fo-20210816-u6sbc72o25gjfveex4mfp2kvi-story.html>.

103 *Id.*

104 See, e.g., OKLA. STAT. ANN. tit. 74, § 150.37 (“[A]ll forensic laboratories . . . shall be ASCLD/LAB accredited.”).

105 NAT’L COMM’N ON FORENSIC SCI., UNIVERSAL ACCREDITATION 1 (2016), <https://www.justice.gov/archives/ncfs/page/file/624026/download>.

106 See ABA CRIMINAL JUSTICE SECTION, *ACHIEVING JUSTICE: FREEING THE INNOCENT, CONVICTING THE GUILTY* 47 (Paul Giannelli & Myrna Raeder eds., 2006).

crime laboratories.¹⁰⁷ This is a good step to ensure that minimal standards exist on paper, but accreditation does not ensure that reliable and consistent work is done in practice. Thus, the American Bar Association called for “demanding written examinations, proficiency testing, continuing education, recertification procedures, an ethical code, and effective disciplinary procedures” for all forensic analysts.¹⁰⁸ Scientists and lawyers have called for comprehensive quality controls for laboratories in the criminal legal system, where a person’s life and liberty are at stake, but so far none have been established.

CONCLUSION

The central problem in forensics is that the function has been treated as a law enforcement rather than scientific function. The result has undermined both public safety and fairness. Innocent people have been convicted, and guilty people have gone free. Quality control failures have caused enormous harm to criminal investigations, victims, and the lives of persons wrongly convicted. We should not allow untrained officers to collect and potentially alter or contaminate evidence at crime scenes instead of enlisting trained scientists. We should not allow people to be arrested and incarcerated based on largely unregulated forensic tests that have never been scientifically tested. Law enforcement should neither run crime labs outright nor otherwise direct and influence lab work. These reforms also provide a model for police more broadly. These standards will likely have to come from lawmakers and not from the courts, which have provided little guidance regarding sound evidence collection during criminal investigations. Fortunately, there are good models for policing forensic evidence, including from labs that have made large-scale changes and from clinical laboratories. Independence, accuracy, and oversight of forensic evidence when it is handled by police, medical examiners, and crime laboratory staff have the potential to dramatically improve the quality of justice in the United States.

107 DUROSE ET AL., *supra* note 36.

108 See ABA CRIMINAL JUSTICE SECTION, *ACHIEVING JUSTICE*, *supra* note 106, at 47.