



Open forensic science*

Jason M. Chin ¹, Gianni Ribeiro² and Alicia Rairden³

1. The University of Sydney Law School, Law School Building (F10), Eastern Avenue, Camperdown Campus, The University of Sydney, New South Wales, NSW 2006, Australia; Institute for Globally Distributed Open Research and Education (IGDORE)

2. Gianni Ribeiro School of Psychology, The University of Queensland, St Lucia QLD 4072, Australia

3. Houston Forensic Science Center, Houston, TX 77002, USA

Corresponding author. E-mail: jason.chin@sydney.edu.au

ABSTRACT

The mainstream sciences are experiencing a revolution of methodology. This revolution was inspired, in part, by the realization that a surprising number of findings in the bioscientific literature could not be replicated or reproduced by independent laboratories. In response, scientific norms and practices are rapidly moving towards openness. These reforms promise many enhancements to the scientific process, notably improved efficiency and reliability of findings. Changes are also underway in the forensic. After years of legal-scientific criticism and several reports from peak scientific bodies, efforts are underway to establish the validity of several forensic practices and ensure forensic scientists perform and present their work in a scientifically valid way.

In this article, the authors suggest that open science reforms are distinctively suited to addressing the problems faced by forensic science. Openness comports with legal and criminal justice values, helping ensure expert forensic evidence is more reliable and susceptible to rational evaluation by the trier of fact. In short, open forensic science allows parties in legal proceedings to understand and assess the strength of the case against them, resulting in fairer outcomes. Moreover, several emerging open science initiatives allow for speedier and more collaborative research.

KEYWORDS: neuroscience, reproducibility, open science

* This article benefited greatly from insightful comments provided by the peer-reviewers. Several of the ideas herein were discussed and improved at the annual meeting of the Evidence-based Forensics Initiative (EBFI). We also thank Kirsty Kent for assistance with data collection. Sarah Hamid provided indefatigable editorial and research support for which we are endlessly grateful.

PART I. INTRODUCTION

Science has long been regarded as ‘self-correcting’, given that it is founded on the replication of earlier work. Over the long term, that principle remains true. In the shorter term, however, the checks and balances that once ensured scientific fidelity have been hobbled. This has compromised the ability of today’s researchers to reproduce others’ findings.¹

Over the past several decades, forensic science has faced immense criticism.² This criticism often reduces to the notion that forensic scientific knowledge has not traditionally been produced and presented in a way that allows judges and juries to assess its reliability. As a result, untested—often invalid—‘science’ contributed to many miscarriages of justice.³ Along a similar timeline, but with almost no express recognition of the issues happening in forensics, a scientific revolution has been occurring in the ‘mainstream sciences’.⁴ This revolution—one focused on methodology—responded to the discovery of several peer-reviewed and published findings that appeared to be false or substantially exaggerated.⁵ Metascientists (ie those who use scientific

- ¹ Francis S. Collins & Lawrence A. Tabak, *Policy: NIH plans to enhance reproducibility*, 505(7485) NATURE 612 (2014).
- ² Suzanne Bell et al., *A call for more science in forensic science*, 115(18) PNAS 4541 (2018); Simon A. Cole, *Toward Evidence-Based Evidence: Supporting Forensic Knowledge Claims in the Post-Daubert Era*, 43(2) TULSA L. REV. 263 (2013); Nicole B. Cáñez & Sandra G. Thompson, *Three Transformative Ideals to Build a Better Crime Lab*, 34(4) GA. ST. U. L. REV. 1007 (2018); Itiel E. Dror, *Biases in forensic experts*, 360(6386) SCIENCE 243 (2018); Gary Edmond, *Forensic Science Evidence and the Conditions for Rational (Jury) Evaluation*, 39(1) MELBOURNE U. L. REV. 77 (2015); Keith A. Findley, *Innocents at Risk: Adversary Imbalance, Forensic Science, and the Search for Truth*, 38(3) SETON HALL L. REV. 893 (2008); Brandon L. Garrett & Peter J. Neufeld, *Invalid Forensic Science Testimony and Wrongful Convictions*, 95(1) VA. L. REV. 1 (2009); Jennifer L. Mnookin, *The Courts, the NAS, and the Future of Forensic Science*, 75(4) BROOK. L. REV. 1209 (2010); NATIONAL RESEARCH COUNCIL, *STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD* (2009) [NAS Report]; PRESIDENT’S COUNCIL OF ADVISORS ON SCIENCE AND TECHNOLOGY, *FORENSIC SCIENCE IN CRIMINAL COURT: ENSURING SCIENTIFIC VALIDITY OF FEATURE-COMPARISON METHODS* (2016) [PCAST Report]; Michael D. Risinger et al., *The Daubert/Kumho Implications of Observer Effects in Forensic Science: Hidden Problems of Expectation and Suggestion*, 90(1) CAL. L. REV. 1 (2002); Michael J. Saks & David L. Faigman, *Failed Forensics: How Forensic Science Lost Its Way and How It Might Yet Find It*, 4 ANNU. REV. LAW SOC. SCI. 149 (2008); Michael J. Saks et al., *Forensic bitmark identification: Weak foundations, exaggerated claims*, 3(3) J. LAW BIOSCI. 1 (2016); William C. Thompson, *Painting the target around the matching profile: the Texas sharpshooter fallacy in forensic DNA interpretation*, 8(3) LAW, PROB. & RIS. 257 (2009).
- ³ In Australia, see Rachel Dioso-Villa, *A repository of wrongful convictions in Australia: First steps toward estimating prevalence and causal contributing factors*, 17(2) FLIN. L. J. 163 (2015); In the U.S., see Brandon L. Garrett & Peter J. Neufeld, *Invalid Forensic Science Testimony and Wrongful Convictions*, 95(1) VA. L. REV. 1 (2009); In Canada, see Bruce A. MacFarlane, *Convicting the Innocent: A Triple Failure of the Justice System*, 31(3) MAN. L. J. 403 (2006).
- ⁴ Andrew Gelman, *The competing narratives of scientific revolution*, <https://andrewgelman.com/2018/08/20/competing-narratives-scientific-revolution/> (accessed 2019); NATIONAL ACADEMIES OF SCIENCES, ENGINEERING, AND MEDICINE, *OPEN SCIENCE BY DESIGN: REALIZING A VISION FOR 21ST CENTURY RESEARCH* (2018) [NASEM Report]; Barbara A. Spellman, *A Short (Personal) Future History of Revolution 2.0*, 10(6) PERSPECT. PSYCHOL. SCI. 886 (2015). In this article, for the purpose of readability, we will draw an admittedly broad distinction between what we will label the ‘mainstream sciences’ and forensic science. The mainstream sciences, such as those reviewed by the National Academics of Sciences, Engineering, and Medicine in its open science report, typically have longer histories, well-established norms, theoretical underpinnings, and are often taught and researched in universities. As we will discuss, forensic science diverges in many ways. Saks & Faigman, *supra* note 2, at 151–152 draw a similar distinction between ‘mainstream’ and forensic science.
- ⁵ NASEM Report, *supra* note 2 at 31–32; Collins & Tabak, *supra* note 1; Leif D. Nelson et al., *Psychology’s Renaissance*, 69 ANNU. REV. PSYCHOL. 511 (2018); Another main motivation behind open science is to make the products of scientific inquiry open to the public, see the discussion *infra* pp. 36–37.

methodology to study the scientific enterprise itself) at the heart of this revolution prescribe more open and transparent methods.⁶ In this article, we evaluate the openness of forensic science in light of the reforms underway in the mainstream sciences. We then consider the distinctive challenges and advantages that openness presents to forensic science, and propose several tangible ways to improve forensic science through open science.

The most authoritative expression (to date) of open science's methods and values is a 2018 Consensus Study Report of the National Academy of Sciences, Engineering, and Medicine ('NASEM Report').⁷ The report reviews the state of openness across several scientific fields and provides a vision for widespread adoption of open science methods. In doing so, it broadly accepts openness as a better way to conduct research:⁸

The overarching principle of open science by design is that research conducted openly and transparently leads to *better science*. Claims are more likely to be credible – or found wanting – when they can be reviewed, critiqued, extended, and reproduced by others.

Similarly, we believe there are at least three pressing reasons for forensic science to adopt a variety of open scientific practices. First, as the NASEM Report notes in the preceding quote, open science enables more thorough analysis of factual claims. Conversely, when science is not conducted transparently, recent metascientific research has found that results are misleading, with actual false positive rates well above what is reported.⁹ This may be acceptable (but not salutary) in the mainstream sciences as the literature may self-correct over time, but it can produce vast injustice in the criminal law context.

Second, and flowing from the first, open and transparent knowledge-generation processes comport with and progress legal values like the presumption of innocence and access to justice. For example, there is well-established imbalance in the state's ability to develop a forensic scientific case against an accused and the accused's ability to assess that case and amass his or her own evidence.¹⁰ This inequity is heightened when the foundational science behind the state's case was conducted opaquely and published in paywalled journals (and then applied in crime labs, which have been described as 'organizational black boxes').¹¹ Similarly, commentators have studied access to justice in terms of *legal* assistance and access to databases of *legal* decisions.¹² However, the *factual* basis of access to justice has largely been neglected. This is unfortunate because

⁶ Simine Vazire, *Implications of the Credibility Revolution for Productivity, Creativity, and Progress*, 13(4) PERSPECT. PSYCHOL. SCI. 411 (2018).

⁷ NASEM Report, *supra* note 2.

⁸ NASEM Report, *supra* note 2 at 107 [*emphasis in original*].

⁹ Joseph P. Simmons et al., *False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant*, 22(11) PSYCHOL. SCI. 1359 (2011); Joseph P. Simmons et al., *False-Positive Citations*, 13(2) PERSPECT. PSYCHOL. SCI. 255 (2018).

¹⁰ Gary Edmond and Kent Roach, *A Contextual Approach to the Admissibility of the State's Forensic Science and Medical Evidence*, 61(3) U. TORONTO L. J. 343, 362 (2011); Findley, *supra* note 2; NAS Report, *supra* note 2 at 11.

¹¹ Cáñez & Thompson, *supra* note 2 at 1007.

¹² Deborah L. Rhode, *Access to Justice*, 69 FORDHAM L. REV. 1785 (2001); John Zeleznikow, *Using Web-Based Legal Decision Support Systems to Improve Access to Justice*, 11(1) ICTL 15 (2010).

if the science behind a case was transparently reported and more affordable to assess, impecunious parties may stand a better chance at mounting a defense.

Third, open science provides a set of tools that may make forensic science more efficient.¹³ Many forensic disciplines have a long way to go in validating their subjective methodologies and in developing objective methodologies.¹⁴ Resource limitations are often severe.¹⁵ To counter such restrictions in the mainstream sciences, open science reformists are developing web platforms and best practices for collaboration and sharing of data and methods.

In the following Part, we briefly review forensic science and the challenges it is currently facing. Part III then delves into the open science movement afoot in the mainstream sciences and assesses forensic science's current level of openness. Next, in Part IV, we tie together the forgoing, examining the ways in which open science is distinctively suited to help improve forensic science. We also address the challenges that forensic science will face in adopting a more open model. Part V concludes.

PART II. THE STATE OF FORENSIC SCIENCE

Forensic science's shortcomings are well-documented, so this section will provide only a brief review with a focus on the areas that may be enhanced through open science reforms.¹⁶ From the beginning, many forensic scientific practices—especially those based on feature comparison—had no basis in academic science.¹⁷ Rather, they arose *ad hoc* in criminal investigations, their development driven by the investigators themselves (rather than independent bodies with scientific training).¹⁸ This meant that many forensic practices developed in a manner that was substantially divorced from scientific structures like the empirical testing of claims, blinding, randomization, and measuring error.¹⁹ Fingerprint identification, for instance, has appeared in U.S. courts since 1911.²⁰ Examiners regularly made identifications about the source of a fingerprint as against all the world and expressly stated that their practice was infallible.²¹ Only in the past two decades have appropriately designed studies been performed and published, supporting the validity of the practice.²²

This nonscientific character of forensic science drew scathing criticism from attentive legal and scientific scholars.²³ Still, courts remained deferential to forensic

¹³ See discussion *infra* at pp. 37–41.

¹⁴ *Infra* at pp. 9–15.

¹⁵ *Ibid.*

¹⁶ See sources at note 2. For a recent review, suggesting there is room for both optimism and pessimism about the state of forensic science, see Jennifer L. Mnookin, *The Uncertain Future of Forensic Science*, 147 DAEDALUS 99 (2018).

¹⁷ PCAST Report, *supra* note 2 at 1 defines feature comparison as: 'methods that attempt to determine whether an evidentiary sample (eg from a crime scene) is or is not associated with a potential "source" sample (eg from a suspect), based on the presence of similar patterns, impressions, or other features in the sample and the source'.

¹⁸ Saks & Faigman, *supra* note 2.

¹⁹ *Id.*

²⁰ See SIMON A. COLE, SUSPECT IDENTITIES: A HISTORY OF FINGERPRINTING AND CRIMINAL IDENTIFICATION (2002).

²¹ *Id.*; Mnookin, *supra* note 2.

²² For earlier research, see Christophe Champod & Ian W. Evett, *A Probabilistic Approach to Fingerprint Evidence*, 51(2), J. FORENSIC IDENTIF. 101 (2001); See generally PCAST Report, *supra* note 2 at 101.

²³ See sources at note 2.

witnesses. There are many reasons for this. For one, legal actors lack the scientific training to appropriately question forensic practices.²⁴ Moreover, forensic scientists have historically been associated with the police and prosecution, making it difficult for the criminally accused to find an independent expert, let alone pay for one.²⁵ With regard to legal structures and safeguards, foundational concepts like *stare decisis* provide little assistance when invalid evidence has historically been admitted into court. Changing the legal standard for admitting evidence from one that defers to the scientific community to one that requires that trial judges engage with scientific concepts seems to have made little difference.²⁶

Perhaps not surprisingly, the widespread admission of untested, invalid, or misleading forensic evidence has contributed to several wrongful convictions.²⁷ Many of these convictions came to light due to the rise of DNA analysis, one of the few forensic sciences to emerge from the mainstream sciences and withstand thorough validation testing.²⁸

Acknowledgement of these wrongful convictions inspired a great deal of research, but none was as momentous as a 2009 report drafted by a National Research Council committee of the National Academy of Sciences (the ‘NAS Report’).²⁹ The report, confirming longstanding worries, catalogued a host of problems:

- deficient training and education among forensic scientists;³⁰
- lack of peer-reviewed and published foundational research establishing the validity of forensic methods;³¹
- lack of protocols to minimize cognitive bias;³²
- insufficient standards for reporting findings and giving testimony;³³ and
- scarce funding to support improvements to any of the foregoing.³⁴

Two federal bodies were created as a result of the NAS Report, but neither has proven as effective as the Committee wished. The Report called for an independent central regulatory body for the forensic sciences.³⁵ While that did not come to be, the U.S. Department of Justice (DOJ) eventually formed the National Commission on Forensic Science (NCFS), an advisory body aimed at providing policy recommendations to the Attorney General.³⁶ Some progress was also made with the DOJ adopting the

²⁴ Saks & Faigman, *supra* note 2, at 161-165.

²⁵ Findley, *supra* note 2.

²⁶ See Jason M. Chin & D’Arcy White, *Forensic Bitemark Identification Evidence in Canada*, 52(1) UBC L. REV. 57 (2019); Peter J. Neufeld, *The (Near) Irrelevance of Daubert to Criminal Justice and Some Suggestions for Reform*, 95(1) 107 AJP (2005); NAS Report, *supra* note 2 at 11; Michael D. Risinger, *Navigating Expert Reliability: Are Criminal Standards of Certainty Being Left on the Dock?*, 64(1) ALB. L. REV. 99 (2000).

²⁷ See sources at note 3 above.

²⁸ NAS Report, *supra* note 2 at 42-44.

²⁹ NAS Report, *supra* note 2.

³⁰ *Id.* at 237-239.

³¹ *Id.* at 187-188.

³² *Id.* at 184-185.

³³ *Id.* at 185-186.

³⁴ *Id.* at 77-83.

³⁵ *Id.* at 81-83.

³⁶ PCAST Report, *supra* note 2 at 22.

NCFS's first recommendations regarding accreditation.³⁷ Just four years after its formation, however, the NCFS was abruptly decommissioned under the new presidential regime.

As recommended by the NAS Report, the National Institute of Standards and Technology (NIST) also took on some new responsibilities. In particular, it created the Organization of Scientific Area Committees (OSAC), which oversees several committees and subcommittees that create and maintain standards for the forensic scientific disciplines. Only time will tell how effective standard-setting can be in regulating forensic science and, more generally, how effective OSAC—the last institutional vestige of the NAS Report—can be.³⁸ The main limitation of OSAC will likely be that it has no express powers to enforce standards and thus they may operate as mere recommendations.

While the shuttering of the NCFS was undoubtedly a setback for those concerned about the state of forensic science, there is reason to think that forensic science finds itself at an inflection point.³⁹ Notably, the NAS Report and a similar 2016 report of the U.S. President's Council of Advisors on Science and Technology (the 'PCAST Report', more on this below) drew attention to longstanding problems in the forensic sciences. The reports appear to have (re)invigorated academic scientific efforts aimed at the forensic sciences.⁴⁰ These efforts are girded by the fact that funding bodies continue to support research in the forensic sciences (although certainly not to the extent many would prefer).⁴¹ Moreover, the media continues to be interested in forensic science-driven controversies.⁴²

Usefully, the PCAST Report delineated a clear framework by which to evaluate the state of forensic science practices, and then compared several feature-comparison disciplines against that standard.⁴³ In short, the PCAST Report said that forensic evidence must be both foundationally valid and then applied in a demonstrably valid way.⁴⁴ As we will discuss in Parts III and IV, transparency and openness assist with both mandates.

By foundationally valid, the PCAST Report authors meant that the method must have been empirically tested to demonstrate that it is '*repeatable, reproducible, and*

³⁷ Bell et al., *supra* note 2, at 4544; PCAST Report, *id.* at 36.

³⁸ Simon A. Cole, *Who Will Regulate American Forensic Science*, 48(3) SETON HALL L. REV. 563 (2018); Simon A. Cole, *A Discouraging Omen: A Critical Evaluation of the Approved Uniform Language for Testimony and Reports for the Forensic Latent Print Discipline*, 34(4) GA. ST. U. L. REV. 1103 (2018).

³⁹ As we will discuss below, the authors of the NASEM Report, *supra* note 2 at 149 suggest that the mainstream sciences are at an inflection point in relation to the adoption of open scientific methods.

⁴⁰ Bell et al., *supra* note 2; Dror, *supra* note 2; Brandon L. Garrett & Chris M. Fabricant, *The Myth of the Reliability Test*, 86(4) FORDHAM L. REV. 1559 (2018).

⁴¹ PCAST Report, *supra* note 2 at 36-38; See also UQ News, *Funding to create forensic expertise*, <https://www.uq.edu.au/news/article/2018/06/funding-create-forensic-expertise> (accessed 2019).

⁴² See Center for Integrity in Forensic Sciences, *Centre for Integrity in Forensic Sciences: Reform in Forensic Sciences, Crime Laboratories, and the Courtroom*, <https://cifsjustice.org/#/main> (accessed 2019); *Last Week Tonight with John Oliver*, 'Forensic Science' (HBO television broadcast, 1 Oct., 2017); *Making a Murderer* (Netflix television broadcast, 2015); *The Staircase* (Netflix television broadcast, 2018).

⁴³ Those practices are: DNA analysis, bitemark analysis, latent fingerprint analysis, firearms analysis, footwear analysis, and hair analysis.

⁴⁴ PCAST, *supra* note 2 at 54-56; Note this is consistent with U.S. and Canadian standards for admitting evidence, see *Federal Rules of Evidence*. Rule 702. https://www.uscourts.gov/sites/default/files/evidence-rules-procedure-dec2017_0.pdf (last accessed 2019) and *R v. J (J-L)* SCC 51 (2000); 2 SCR 600 [2000].

accurate, at levels that have been measured and are appropriate to the intended application'.⁴⁵ Repeatable in this formulation, refers to intra-examiner reliability—the same examiner should come to the same result over time. Reproducible refers to inter-examiner reliability—different examiners should come to the same conclusions. And finally, accuracy can be thought of as error control. The method should have a known and tolerable level of error, avoiding false positives and false negatives. Of the feature-comparison disciplines reviewed by PCAST in 2016, only DNA analysis of single-source samples and fingerprint analysis were foundationally valid.⁴⁶

As to applied validity, the PCAST Report explained that it has two components.⁴⁷ First, the examiner must be demonstrably capable of applying the method. This capacity should typically be supported by proficiency tests.⁴⁸ Proficiency tests are empirical demonstrations that the examiner can accurately make the relevant judgment in realistic situations. Ideally, these tests should be inserted into the examiner's coursework such that he or she is unaware that he or she is being tested. Second, the examiner must have faithfully applied that method and reported any uncertainty in the conclusion (eg the false positive rate of the method).⁴⁹

While the exigency of establishing foundational and applied validity may seem obvious to outside observers, the PCAST Report found that substantial hurdles still exist in establishing them across most of forensic science. Indeed, it found that considerable work still needed to be done in all of the feature comparison disciplines it reviewed. For instance, although the report found that fingerprint analysis is foundationally valid, proficiency testing should be improved (eg the tests should be more representative of actual casework), and examiners do not always faithfully apply the method.⁵⁰

Finally, the PCAST Report strongly recommended that forensic science develop more objective methods (ie those generally less reliant on subjective judgment), often through automated image analysis.⁵¹ It did so for several reasons: objective methods are generally more transparent and reliable, and they present a lower risk of human error and bias.⁵² Significant progress has been made towards developing objective systems for fingerprint and firearms analysis. However, a key limitation going forward is the lack of a large body of stimuli (eg a database containing images of fingerprints with a known ground truth). As we will see, open science may provide viable responses to such limitations.

PART III. IMPROVING SCIENCE THROUGH OPENNESS AND TRANSPARENCY

In contrast to forensic science, concepts like validation testing and blinding are orthodox in much of mainstream science. That said, undisclosed flexibility in the scientific process has still allowed for researchers' biases and expectations to influence results.

⁴⁵ PCAST, *id.* at 47 [emphasis in original].

⁴⁶ *Id.* at 75, 101.

⁴⁷ *Id.* at 56.

⁴⁸ *Id.* at 57-58.

⁴⁹ *Id.* at 56.

⁵⁰ *Id.* 102.

⁵¹ *Id.* at 125-126. Automated image analysis seeks to train algorithms to determine if two images came from the same source.

⁵² *Id.*

Recent metascientific research has explored how to improve the scientific process, often recommending openness and transparency that would dissuade and reveal more subtle forms of researcher bias. Predating many of these concerns, scholars and advocacy groups have long campaigned for more open access to the products of research.⁵³ They have noted that research data and findings often sit behind paywalls, making it difficult for those without institutional access to use and verify that knowledge. The open science movement, therefore, encompasses the aims of democratizing knowledge and producing knowledge that is more trustworthy. As we will see, both aims are fundamentally important to forensic science's ongoing development.⁵⁴

In this Part, after a brief review of recent controversies in mainstream science, we will introduce the open science movement and its more specific manifestations (eg open data, open access journals). As part of this discussion, we will assess the degree to which forensic science is adopting these reforms. We generally find that while there have been some promising developments, there is still much work left to do in opening forensic science (but many reasons to take on that work).

Science in Crisis

Concerns about the number of scientific findings that may be false or exaggerated have percolated for years, but have reached a fever pitch in the past decade or so.⁵⁵ For instance, in 2016, the journal *Nature* asked approximately 1500 scientists whether science was experiencing a reproducibility 'crisis'.⁵⁶ The researchers found that 52% of those surveyed believed there was a significant crisis, 38% believed there was a slight crisis, and only 3% thought there was no crisis.

Note that there is some ambiguity and inconsistency in the literature in the use of the terms 'reproducibility' and 'replicability'.⁵⁷ For the purposes of this article, we will define replication as repeating a study exactly with new data to determine if the same result is achieved. We will refer to reproduction as repeating the analysis used by an existing study on its own data to see if the results are the same. Both replicability and reproducibility are thwarted when published reports do not provide enough information about how the study was conducted or provide the raw data for re-analysis.⁵⁸ But,

⁵³ NASEM Report, *supra* note 2 at 23-58; *Paywall: The Business of Scholarship* (Online documentary, 2018) <https://paywallthemovie.com> (last accessed 2019).

⁵⁴ There is certainly much more to say about the epistemology of forensic science and its relationship with what we have termed the mainstream sciences. This article, however, focuses more pragmatically on the benefits and challenges that will go along with forensic science becoming more open and transparently. See Saks & Faigman, *supra* note 2 for a broader discussion of the relationship between mainstream and forensic science.

⁵⁵ For our definitions of reproducibility and replicability, see *infra* p. 17.

⁵⁶ Monya Baker, *1,500 scientists lift the lid on reproducibility*, 533 NATURE 452 (2016). The survey defined reproducibility as follows: 'For the purposes of this survey, we consider a study to be reproduced when its findings are confirmed in similar experimental systems (these may include slight variations in methods or materials.) By contrast, a study is replicated when it is repeated exactly, using the same reagents. This survey talks about the larger issue of reproducibility of results, not just replication'.

⁵⁷ See Leonard P. Freedman et al., *The Economics of Reproducibility in Preclinical Research*, 16(4) PLOS BIOLOGY 1, at 2 (2015); Victoria Stodden, *Enhancing reproducibility for computational methods*, 354(6317) SCIENCE 1240 (2016). Compare the definitions used in these works with the definition used in the *Nature* study, *Id.*

⁵⁸ Jelte M. Wicherts et al., *Willingness to Share Research Data is Related to the Strength of the Evidence and the Quality of Reporting of Statistical Results*, 6(11) 1 PLOS ONE (2011); *Reducing our irreproducibility*, 496 NATURE 398 (2013): 'The problems arise in laboratories, but journals such as this one compound them when they fail to exert sufficient scrutiny over the results that they publish, and when they do not publish enough

more than studies not providing enough information to replicate them or reproduce their findings—in many cases, when replication attempts *have* been conducted, the results have contradicted the original findings.⁵⁹

Indeed, The *Nature* survey was released in the wake of several large-scale failures of replication. For example, in social science, one of the largest efforts to date attempted to replicate 100 studies published in three of psychology's top journals.⁶⁰ The researchers found the same result with the same level of statistical certainty in approximately one third of the studies and generally found considerably smaller effect sizes. In 2018, another collaboration of researchers attempted to replicate the findings of 21 social scientific studies published in *Nature* and *Science*.⁶¹ They found the originally reported effect in 13 of the attempts and, overall, effect sizes were about 50% smaller than in the original studies.

In medicine, one influential review of preclinical research found reports of irreproducibility and irreproducibility ranging from 89%–51%.⁶² Using a conservative estimate, this corresponds with \$28B of lost research funds. Findings like these led Francis Collins (Director of the U.S. National Institutes of Health) and Lawrence Tabak to state that the checks and balances of science have been 'hobbled'. Troublingly, human trials have also not escaped criticism. Researchers in the UK, for instance, found in 2018 that approximately 50% of studies were in contravention of EU laws requiring reporting of results (similar laws exist in the US).⁶³ And, studies replicating clinical studies also often show contradictory results or significantly smaller effects.⁶⁴

While we have focused on social science and medicine, that is largely because these fields have been unusually proactive in examining their own practices and admitting

information for other researchers to assess results properly.'; In other words, studies have long been published with methodology sections, but they have not always been detailed enough to allow for other researchers to fully scrutinize them, and, if desired, replicate them. On the other hand, publishing any raw data is not the norm in many fields and never has been.

⁵⁹ For a review, see Nelson et al., *supra* note 5, at 17.3–17.4 (2018); Jacob S. Sherkow, *Patent Law's Reproducibility Paradox*, 66 DUKE L.J. 845, at 852–865 (2017).

⁶⁰ Open Science Collaboration, *Estimating the reproducibility of psychological science*, 349(6251) SCIENCE 943 (2016); For a criticism of this, see Daniel T. Gilbert et al., *Comment on 'Estimating the reproducibility of psychological science'*, 351(6277) SCIENCE 1037 (2016). For other largescale replication efforts in psychology, see: Richard A. Klein et al., *Investigating Variation in Replicability: A 'Many Labs' Replication Project*, 45(3) SOC. PSYCHOL. 142 (2014) [Many Labs 1], Richard A. Klein et al., *Many Labs 2: Investigating Variation in Replicability Across Sample and Setting*, 1(4) AMPPS 443 (2018) [Many Labs 2], and Charles R. Ebersole et al., *Many Labs 3: Evaluating participant pool quality across the academic semester via replication*, 67 J. EXP. SOC. PSYCHOL. 68 (2016) [Many Labs 3].

⁶¹ Colin F. Camerer et al., *Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015*, 2 NAT. HUM. BEHAV. 637 (2018).

⁶² Freedman et al., *supra* note 57, at 2; See also C. Glenn Begley and Lee M. Ellis, *Raise standards for preclinical cancer research*, 483 NATURE 531 (2012), which found that, in its attempt to replicate 53 landmark studies, only six were successfully replicable.

⁶³ Ben Goldacre, *Compliance with requirement to report results on the EU Clinical Trials Register: cohort study and web resource*, 362 BMJ 1 (2018).

⁶⁴ John P. A. Ioannidis, *Contradicted and Initially Stronger Effects in Highly Cited Clinical Research*, 294(2) JAMA 218 (2005).

deficiencies.⁶⁵ Indeed, similar problems have been reported in a variety of fields.⁶⁶ In Neuroscientific research, for instance, many of the correlations reported between brain activation and behavior or personality measures are far higher than is statistically possible.⁶⁷ Further, an analysis of over 3000 papers in the cognitive neuroscience field were underpowered to detect true effects,⁶⁸ suggesting a false discovery rate of over 50% across the discipline.⁶⁹

Contributors to the Crisis

Critically, the controversies recounted above have been followed by a raft of metascientific research aimed at determining why so many studies are proving difficult to replicate and reproduce. Much of this work builds on historic concerns about the research process, lending such concerns support from modern quantitative methods.⁷⁰ We will now briefly review a selection of the culprits identified by recent metascientific study: 'questionable research practices' (QRPs), 'publication bias', 'spin', lack of replication, small sample sizes, and overreliance on simplistic statistical methods.⁷¹

QRPs exploit flexibility in research methods and reporting practices to make a researcher's results seem more persuasive than they actually are.⁷² Such practices include deciding whether to exclude observations after looking at how this would affect the overall results, measuring a phenomenon several different ways but only disclosing those measures that support the hypothesis and strategically stopping data collection when results reach some level of statistical confidence.⁷³ These tactics rest in

⁶⁵ See NASEM Report, *supra* note 2 at 1; Christie Aschwanden, *Psychology's Replication Crisis Made the Field Better*, <https://fivethirtyeight.com/features/psychologys-replication-crisis-has-made-the-field-better/> (accessed 2019).

⁶⁶ See Richard Border et al, *No Support for Historical Candidate Gene or Candidate Gene-by-Interaction Hypotheses for Major Depression Across Multiple Large Samples*, AM. J. PSYCHIATRY (forthcoming 2019); ANDREW C. CHANG & PHILLIP LI, IS ECONOMICS RESEARCH REPLICABLE? SIXTY PUBLISHED PAPERS FROM THIRTEEN JOURNALS SAY 'USUALLY NOT' (Board of Governors of the Federal Reserve System, Finance and Economics Discussion Series Paper 2015-083); Hannah Fraser et al., *Questionable research practices in ecology and evolution*, 13(7) PLOS ONE 1 (2018); Daiping Wang et al, *Irreproducible text-book 'knowledge': The effects of color bands on zebra finch fitness*, 72(4) EVOLUTION 961 (2018).

⁶⁷ Edward Vul et al., *Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition*, 4(3) PERSPECT. SOC. PSYCHOL. 274 (2009); Anders Eklund et al., *Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates*, 113(28) PNAS 7900 (2016).

⁶⁸ For a definition of power, see the text accompanying *infra* note 85.

⁶⁹ Denes Szucs & John P. A. Ioannidis, *Empirical assessment of published effect sizes and power in the recent cognitive neuroscience and psychology literature*, 15(3) PLOS BIOLOGY 1 (2017). See also Chuan-Peng Hu et al., *Open science as a better gatekeeper for science and society: a perspective from neurolaw*, 63 SCI. BULL. 1529 (2018).

⁷⁰ See Spellman, *supra* note 4.

⁷¹ For the purposes of this review, we focus on methods that have not historically been condemned, but still introduced error into the literature. That said, research fabrication and fraud has certainly contributed to irreproducibility. Best estimates for fabrication suggest that about 1% of researchers have engaged in it at some point. See Danielle Fanelli, *How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data*, 4(5) PLOS ONE 1 (2009).

⁷² Leslie K. John et al., *Measuring the Prevalence of Questionable Research Practices With Incentives for Truth Telling*, 23(5) PSYCHOL. SCI. 524 (2012): 'Although cases of overt scientific misconduct have received significant media attention recently, exploitation of the gray area of acceptable practice is certainly much more prevalent, and may be more damaging to the academic enterprise in the long run, than outright fraud'.

⁷³ *Id.*

a gray area of scientific practice that many once viewed as defensible, trivial, or even normative.⁷⁴

Views that would minimize the harmful impact of QRPs are now demonstrably untenable. In a widely influential paper,⁷⁵ Joseph Simmons and colleagues employed a quantitative simulation to demonstrate that the use of QRPs increased the actual false positive rate of a literature well-beyond its reported false positive rate.⁷⁶ Use of four QRPs increased a notionally 5% false positive rate to approximately 60%. Troublingly, metascientific research has also found that QRPs are, in fact, widely used. Anonymous surveys in psychology, ecology, and evolutionary biology find self-reported usage of QRPs ranging from approximately 3% to 60%, depending on the QRP in question.⁷⁷ Note, however, that recent empirical work in psychology suggests that initial estimates of QRP use based on survey studies are inflated.⁷⁸

QRPs conspire with ‘publication bias’ and ‘spin’ to provide, in some cases, a deeply misleading view of a research literature. The term publication bias refers to published scientific literature containing systematic biases as to what types of articles are published. This includes the empirically-founded observation that studies that find no effect (eg a drug had no discernable impact on a disease) tend to be published much less frequently than studies that find some effect (ie the null studies languish in file drawers, hence the colloquialism, the ‘file-drawer effect’).⁷⁹ This can contribute to strategic research patterns. Researchers may, for instance, employ the tactic of performing several underpowered studies (ie they collect few observations) that vary in small ways, and only reporting those studies that find positive effects.⁸⁰

Another instance of publication bias is a preference for novel results. Journals typically prefer to publish articles that purport to show some heretofore undiscovered phenomenon, rather than studies that simply attempt to replicate previous studies.⁸¹ This is problematic because, as mentioned above, replication lends credibility to previous research and can help uncover spurious findings.

Even if publication bias is overcome, published negative results are cited less frequently (ie citation bias).⁸² Further, published reports may be ‘spun’.⁸³ In other words, the report may suggest that some positive effect exists by emphasizing positive findings

⁷⁴ *Id.*

⁷⁵ Simmons et al., *False Positive Citations*, *supra* note 9.

⁷⁶ Simmons et al., *False-Positive Psychology*, *supra* note 9.

⁷⁷ Fraser et al., *supra* note 65; John et al., *supra* note 72.

⁷⁸ Klaus Fiedler & Norbert Schwarz, *Questionable Research Practices Revisited*, 7(1) SOC. PSYCHOL. PERS. SCI 45 (2015).

⁷⁹ Anthony G. Greenwald, *Consequences of Prejudice Against the Null Hypothesis*, 82(1) PSYCHOL. BULL. 1 (1975); Robert Rosenthal, *The file drawer problem and tolerance for null results*, 86(3) PSYCHOL. BULL. 638 (1979).

⁸⁰ Marjan Bakker et al., *The Rules of the Game Called Psychological Science*, 7(6) PERSPECT. PSYCHOL. SCI. 543 (2012); Matthew C. Makel et al., *Replications in Psychology Research: How Often Do They Really Occur?*, 7(6) PERSPECT. PSYCHOL. SCI. 537 (2012);

⁸¹ James W. Neuliep & Rick Crandall, *Reviewer bias against replication research*, 8(6) SBP JOURNAL 21 (1993).

⁸² Bram Duyx et al., *Scientific citations favor positive results: a systematic review and metaanalysis*, 88 J. CLIN. EPIDEMIOL. 92 (2017).

⁸³ Isabelle Boutron et al., *Reporting and Interpretation of Randomized Controlled Trials With Statistically Non-significant Results for Primary Outcomes*, 303(20) JAMA 2058 (2010).

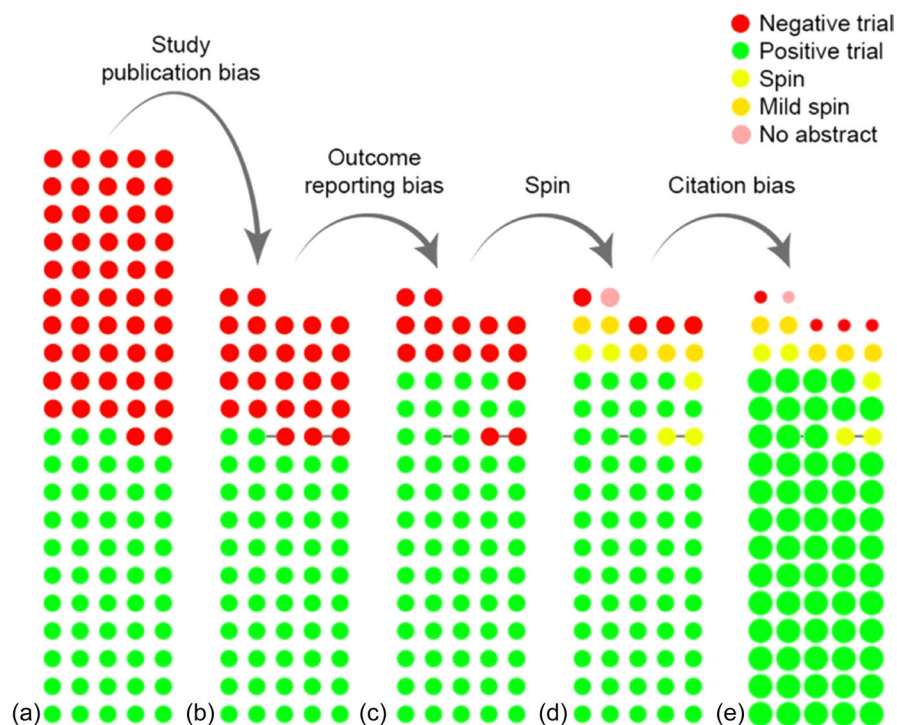


Figure 1. This figure demonstrates the compounding effects of publication bias, reporting bias, spin, and citation bias in research on a treatment for depression. The confluence of these forms of bias results in a very misleading picture of the treatment's efficacy. Looking at *all of the studies*, only 50% found the treatment effective. The combined effect of the biases studied by these researchers makes it appear like the vast majority of the studies demonstrated that the treatment was effective (Reused under an unrestricted Creative Commons Attribution license. Original authors: de Vries et al., *supra* note 84.).

and deemphasizing negative ones (although both findings can be found by a careful reader, distinguishing this from some QRPs that would suppress the negative finding altogether).⁸⁴

Y.A. de Vries and colleagues recently collected and analysed 105 trials of antidepressant medication (see Figure 1). They found that a confluence of QRPs, publication bias, citation bias, and spin produced a deeply misleading portrait of a body of knowledge. In the literature they researched, it appeared (to a reader attending to the primary message of the published reports) that only a small portion of the studies found a negative effect: the drug appeared effective. However, once the researchers unearthed the unpublished studies and waded through spin and citation bias, *half of the studies were actually negative*. Importantly, de Vries and colleagues could only do this because clinical trials are regulated such that it is possible to find true number studies that are being performed (ie they are registered, see below). As we will see below, the same cannot be said for forensic science.

⁸⁴ Y.A. de Vries et al., *The cumulative effect of reporting and citation biases on the apparent efficacy of treatments: the case of depression*, 48(15) PSYCHOL. MED. 2453 (2018).

Finally, small sample sizes and overly simplistic statistical methods have contributed to the replicability crisis. During the opening salvos of the crisis, John Ioannidis famously predicted that over half of published findings were false (based on a theoretical model) because, among other reasons, studies typically do not use large enough samples to find the effects they are looking for (ie they are underpowered).⁸⁵ Moreover, the commonly used statistical method of null hypothesis significance testing (NHST) is often applied with little thought.⁸⁶ QRPs, as mentioned above, can render the results of NHST misleading by producing false positive rates that underestimate the true false positive rate. And unlike other statistical methods (eg Bayesian), NHST does not take into account the *a priori* likelihood of the hypothesis.⁸⁷

The Open Science Response (and Forensic Science's Place Within it)

The scientific community is rapidly adopting transparency-related reforms as a way to improve science.⁸⁸ The NASEM Report, for instance, strongly endorsed many open science reforms: 'open science strengthens the self-correcting mechanisms inherent in the research enterprise'.⁸⁹ In this section, we review some these reforms. It is important to note, however, that open science should not be construed as a panacea for all flaws and inefficiencies in the scientific process.⁹⁰ In both the mainstream sciences and forensic science, change depends on the concerted efforts of: (1) oversight and funding bodies; (2) journal editors and publishers; and, (3) the researchers themselves. We will review the roles of these stakeholders before delving into the specifics of open science reform in the forensic sciences.

Drivers of Change in Mainstream and Forensic Science

As to leadership, the National Institute of Standards and Technology (NIST) may be best placed to guide the move to open forensic science. Indeed, both the NASEM Report and the PCAST Report identified NIST as a crucial leader in the movements they described.⁹¹ Within the forensic sciences, NIST's tasks may include periodically reviewing the state of foundational validity in various disciplines, advising on the design and execution of validation studies, creating and disseminating datasets, and providing grant support. These jobs may be guided by NIST's broader role as a leader in the open science movement as it encourages open practices and, in some cases, makes funding contingent on them. Further, the National Science Foundation—active in funding

⁸⁵ John P. A. Ioannidis, *Why Most Published Research Findings Are False*, 2(8) PLOS MED. 696 (2005).

⁸⁶ Daniel J. Benjamin et al., *Redefine statistical significance*, 2(1) NAT. HUM. BEHAV. 6 (2018); Regina Nuzzo, *Scientific method: Statistical errors*, 506 NATURE 150 (2014): 'P values, the "gold standard" of statistical validity, are not as reliable as many scientists assume'. Several other criticisms have been levied against NHST as traditionally performed (eg, reported results often do not include effect sizes and confidence intervals, which are useful for providing a full understanding of a research finding), see: Geoff Cumming, *The new statistics: why and how*, 25(1) PSYCHOL. SCI. 7 (2013). See also Geoff Cumming et al., *Statistical reform in psychology: Is anything changing?*, 18(3) PSYCHOL. SCI. (2007) 230.

⁸⁷ John K. Kruschke & Torrin M. Liddell, *The Bayesian New Statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective*, 25(1) PSYCHON. BULL. REV. 178 (2018).

⁸⁸ Marcus R. Munafò et al., *A manifesto for reproducible science*, 1(1) NAT. HUM. BEHAV. 1 (2017).

⁸⁹ NASEM Report, *supra* note 2 at 32.

⁹⁰ *Id.*: 'Yet open science is not the only factor or solution to addressing the reproducibility issue, and open science will not automatically solve whatever problems there are.'

⁹¹ *Id.* at 145-146; PCAST Report, *supra* note 2 at 14-15, 124-126.

both the mainstream and forensic sciences—is already requiring that researchers engage in open practices.⁹²

Journal editors and publishers will also be instrumental in the transition to open forensic science, as they are in the mainstream sciences.⁹³ This will especially be so if forensic science moves in the direction urged by the PCAST Report, with forensic scientists adopting an increased interest in publishing their work in reputable journals.⁹⁴ Currently, one of the most influential models for openness in peer-review and publishing is the Transparency and Openness Promotion (TOP) guidelines.

The TOP guidelines, first published in *Science* in 2015, are a standard set of guidelines for transparency and reproducibility practices across journals.⁹⁵ They are comprised of eight standards: citation (eg citing data, materials, and code), data transparency (eg posting data to an online database), analytic methods (code) transparency, research materials transparency (eg surveys and stimuli), study preregistration, analysis plan preregistration (see our discussion of preregistration below), and replication.⁹⁶ The TOP Committee defined three levels for each standard which range from journals simply encouraging the standards (level 0), to requiring and verifying that the articles have met each standard (3).⁹⁷ The *Center for Open Science* provides tools and guidance for organizations wishing to implement TOP and keeps a list of those which have agreed to consider the TOP guidelines (signatories) and those which have implemented them at some level. As of March 2019, over 5000 journals and organizations are signatories and 1000 have implemented them.⁹⁸ A number of journals have also adopted a badge system to acknowledge papers that are preregistered and have open data and open materials.⁹⁹ This initiative is promising, with open data in a leading psychological journal increasing from 3%–23% after implementation of the badge system.¹⁰⁰

Beyond government organizations and publishers, adoption of open science in forensics will depend on individual researchers and practitioners. This is already

⁹² NASEM Report, *id.* at 89; PCAST Report, *id.* at 36–38; See also Munafo et al., *supra* note 88, at 7.

⁹³ NASEM Report, *id.* at 129–130.

⁹⁴ PCAST Report, *supra* note 2 at 11, 125: ‘Finally, we believe that the state of forensic science would be improved if papers on the foundational validity of forensic feature-comparison methods were published in leading scientific journals rather than in forensic-science journals, where, owing to weaknesses in the research culture of the forensic science community discussed in this report, the standards for peer review are less rigorous.’

⁹⁵ Brian A. Nosek et al., *Promoting an open research culture*, 348(6242) *SCIENCE* 1422 (2015); See also the Consolidated Standards of Reporting Trials (CONSORT) statement, which ‘provides guidance for clear, complete, and accurate reporting of randomized control trials.’ In Munafo et al., *supra* note 88, at 4.

⁹⁶ Center for Open Science, Guidelines for Transparency and Openness Promotion (TOP) in Journal Policies and Practices: ‘The Top Guidelines’ Version 1.0.1, https://osf.io/ud578/?_ga=2.131977612.2018415677.1536528864- (accessed Sep. 9, 2018). We define preregistration at *infra* pp 28–32.

⁹⁷ For example, consider the various levels of the ‘data’ transparency guideline. Level 0 corresponds to no journal requirements or mere encouragement. From there, 1 corresponds to requiring authors state whether data is available, 2 is requiring data be posted to a trusted online repository with limited exceptions, and 3 is editorial verification of the data analysis of that data. For A summary of levels for all TOP, see Center for Open Science, *Top Guidelines: The Standards*, <https://cos.io/our-services/top-guidelines/> (accessed 2019).

⁹⁸ For a full list of journals who are signatories to the TOP Guidelines, see Center for Open Science, *Top Guidelines: Current Signatories*, <https://cos.io/our-services/top-guidelines/> (accessed 2019).

⁹⁹ Center for Open Science, *Badges to Acknowledge Open Practices*, <https://osf.io/tvyxz/> (accessed Sep. 8, 2018); Eric Eich, *Business Not as Usual*, 25(1) *PSYCHOL. SCI.* 3 (2014).

¹⁰⁰ Mallory C. Kidwell et al., *Badges to Acknowledge Open Practices: A Simple, Low-Cost, Effective Method for Increasing Transparency*, 14(5) *PLOS BIOL.* 1 (2016).

beginning. Among practitioners, the Netherlands Forensics Institute (NFI) is adopting strong transparency reforms with respect to any quality-control related issues in their labs.¹⁰¹ Further, forensics researchers may use online tools like the Open Science Framework (OSF).¹⁰² The OSF is a free online platform for open science where any researcher (with an academic affiliation or not) can create a webpage for a research project and use that to share data, analysis, and materials. It also contains several tools for collaboration. An example (albeit a rare one) of the use of the OSF in forensic science research is a recent Australian state police-federal government funded collaboration between a university cognitive science lab, which has adopted open science reforms, and several local police services.¹⁰³

Through the remainder of this Part, it may be instructive to compare specific openness-related reforms in the mainstream sciences to the state of openness in forensic scientific research (Table 1).¹⁰⁴ To that end, we performed some preliminary research about openness in forensic science by reviewing the policies of forensic scientific journals (see Appendix A for a description of our search). We identified 30 forensic science journals and recorded whether they were open access, a TOP signatory, and adopted any of the eight TOP standards. We acknowledge that using journal standards as an index for openness is incomplete, not least because cultural differences between forensic and academic science have produced different values surrounding publishing.¹⁰⁵ Still, as mentioned above, journals can be a major driver of reform and so it is useful to see what is happening among them.

Preregistration and Registered Reports

One of the most important developments emerging from the open science movement is preregistration (ie prespecifying research choices in a way that cannot be changed after seeing the data). During preregistration, researchers specify their research plans prior to carrying out the research. Preregistration puts an emphasis on making methodological and statistical decisions ahead of time: calculating sample sizes, determining data exclusion and stopping rules, making predictions and hypotheses, and establishing data analysis plans (ie which analyses will be performed to test each hypothesis?).¹⁰⁶ Once submitted to an online platform, such as the OSF, or AsPredicted, preregistrations are time-stamped and uneditable.¹⁰⁷ Preregistration is required in some areas of clinical

¹⁰¹ PCAST Report, *supra* note 2 at 74-75. See also the innovations at the Houston Forensic Science Center, Cáñez & Thompson, *supra* note 2.

¹⁰² See *infra* note 107.

¹⁰³ UQNews, *supra* note 41.

¹⁰⁴ For a full spreadsheet (Table 1 only reports select TOP standards due to space constraints), see: Center for Open Science, Openness of Forensic Journals, <https://osf.io/Spk7j/> (accessed 2019).

¹⁰⁵ Simon A. Cole, *Forensic culture as epistemic culture: the sociology of forensic science*, 44(1) STUD. HIST. PHILOS. SCI. C. 36 (2013).

¹⁰⁶ Association for Psychological Science, What is Preregistration, Anyway?, <https://www.psychologicalscience.org/publications/observer/obsonline/what-is-preregistration-anyway.html> (accessed Sep. 7, 2018); Brian A. Nosek et al., *The Preregistration Revolution*, 115(11) PNAS 2600 (2018).

¹⁰⁷ For the Open Science Framework repository platform, see: Center for Open Science, Open Science Framework: A scholarly commons to connect the entire research cycle, <https://osf.io> (accessed 2019). For the AsPredicted platform, see: AsPredicted: Pre-Registration Made Easy, <https://aspredicted.org> (accessed 2019). The OSF provides a helpful guide for what to include in a preregistration, see Center for Open Science, Enter the Preregistration Challenge, <http://help.osf.io/m/registrations/1/546603-enter-the-preregistration-challenge> (accessed Sep. 7, 2018).

Table 1. Forensic science journals, their impact factors, whether they are TOP signatories or open access, and their status on selected TOP standards (0 is top level 0; 0/Enc indicates the journal has not adopted TOP, but still expressly encourages the relevant standard). All journals received a 0 for the five TOP standards omitted from Table 1. See the full table at <https://osf.io/5pk7j/>.

<i>Journal</i>	<i>Impact factor</i>	<i>TOP signatory?</i>	<i>Open access?</i>	<i>TOP citations</i>	<i>TOP data</i>	<i>TOP code</i>
Am. J. of Forensic Medicine and Pathology	.64	No	Hybrid	0	0	0
Aus. J. of Forensic Medicine	.94	No	Hybrid	0/Enc	0/Enc	0
Environmental Forensics	.68	No	Hybrid	0	0/Enc	0
Forensic Chemistry		Yes	Hybrid	0/Enc	0/Enc	0/Enc
Forensic Science International	1.974	Yes	Hybrid	0/Enc	0/Enc	0/Enc
Forensic Science International: Genetics	5.64	Yes	Hybrid	0/Enc	0/Enc	0/Enc
Forensic Science International: Synergy		No	Open	0/Enc	0/Enc	0/Enc
Forensic Science Rev.	2.71	No	Closed	0	0	0
Forensic Science, Medicine, and Pathology	2.03	Yes	Hybrid	0	0/Enc	0
Forensic Toxicology	3.92	Yes	Hybrid	0	0	0
Indian J. of Forensic Medicine and Toxicology	.05	No	Closed	0	0	0
Int. J. of Forensic Science & Pathology	.342	No	Hybrid	0	0	0
Int. J. of Legal Medicine	2.31	No	Hybrid	0	0/Enc	0
J. of Forensic and Legal Medicine	1.10	Yes	Hybrid	0/Enc	0/Enc	0/Enc
J. of Forensic Medicine	0	No	Hybrid	0	0	0

Table 1. (Continued)

<i>Journal</i>	<i>Impact factor</i>	<i>TOP signatory?</i>	<i>Open access?</i>	<i>TOP citations</i>	<i>TOP data</i>	<i>TOP code</i>
J. of Forensic Practice	.59	Yes	Hybrid	0/Enc	0/Enc	0/Enc
J. of Forensic Radiology and Imaging	.51	Yes	Hybrid	0/Enc	0/Enc	0/Enc
J. of Forensic Research	.32	No	Hybrid	0	0	0
J. of Forensic Science & Criminology		No	Hybrid	0	0	0
J. of Forensic Sciences	1.18	No	Hybrid	0	0	0
J. of Forensic Toxicology & Pharmacology	.25	No	Hybrid	0	0	0
J. of Law Medicine and Ethics	.99	No	Hybrid	0	0	0
J. of Medical Toxicology and Clinical Forensic Medicine	0	No	Hybrid	0	0	0
Legal Medicine	1.25	Yes	Hybrid	0/Enc	0/Enc	0/Enc
Medical Law Review	1.10	No	Hybrid	0	0	0
Medicine, Science and the Law	.58	No	Hybrid	0	0	0
Rechtsmedizin (Legal Medicine)	.64	No	Hybrid	0	0	0
Regulatory Toxicology and Pharmacology	2.81	Yes	Hybrid	0/Enc	0/Enc	0/Enc
Romanian Journal of Legal Medicine	.32	No	Closed	0	0	0
Science & Justice	1.85	Yes	Hybrid	0/Enc	0/Enc	0/Enc

medical research.¹⁰⁸ In other fields, it is becoming increasingly popular: in 2012, there were merely 38 preregistrations on the OSF repository, a number that has grown to over 12,000 in 2017.¹⁰⁹

Preregistration helps limit ‘over-interpretation of noise’¹¹⁰ by making any data-contingent analytic choices salient.¹¹¹ In other words, it becomes more difficult to engage in QRPs because researchers can no longer selectively exclude data and measures that run counter to their hypothesis, tactics that would give their findings a superficial gleam of credibility.¹¹² When there are no preset decision rules, it is easy for even highly-trained academic scientists to convince themselves that they would have made the choice regardless of how the data looked:

Once we obtain an unexpected result, we are likely to reconstruct our histories and perceive the outcome as something that we could have, even did, anticipate all along—converting a discovery into a confirmatory result. And even if we resist those reasoning biases in the moment, after a few months, we might simply forget the details, whether we had hypothesized the moderator, had good justification for one set of exclusion criteria compared with another, and had really thought that the one dependent variable that showed a significant effect was the key outcome.¹¹³

Preregistration can also help address publication bias, especially with respect to the failure to publish negative findings or those that do not support a particular research agenda. Indeed, a 2018 study found increased reporting of null (ie negative) findings associated with the rise of preregistration.¹¹⁴

Within forensic science, our search did not uncover any peer-reviewed journal that encourages (or even expressly mentions) preregistration.¹¹⁵ That is not to say, however, that it has been altogether ignored. The PCAST Report stated that validation studies should be preregistered (although the studies they relied on were not preregistered): ‘The study design and analysis framework should be specified in advance. In validation studies, it is inappropriate to modify the protocol afterwards based on the results’.¹¹⁶

We concur with the PCAST Report. In fact, preregistration may be even more important in forensic scientific validation research. There are many analytic choices

¹⁰⁸ National Institute of Health, FDAAA 801 and the Final Rule, <https://clinicaltrials.gov/ct2/manage-recs/fdaaa> (accessed Sep. 9, 2018); But see also Goldacre, *supra* note 63.

¹⁰⁹ Brian A. Nosek & D. Stephan Lindsay, Preregistration Becoming the Norm in Psychological Science, <https://www.psychologicalscience.org/observer/preregistration-becoming-the-norm-in-psychological-science> (accessed Sep. 9, 2018).

¹¹⁰ Munafò et al., *supra* note 88, at 1.

¹¹¹ *Id.* at 3.

¹¹² NASEM Report, *supra* note 2 at 109; Gerard GMH Swaen et al., *False positive outcomes and design characteristics in occupational cancer epidemiology studies*, 30(5) INT. J. EPIDEMIOL. 948 (2001); Anna Elisabeth van’t Veer & Roger Giner-Sorolla, *Pre-Registration in Social Psychology a Discussion and Suggested Template*, 67(1) J. EXP. SOC. PSYCHOL. 2 (2016); Simine Vazire, *supra* note 6.

¹¹³ Brian A. Nosek et al., *Scientific Utopia II: Restructuring Incentives and Practices to Promote Truth over Publishability*, 7(6) PERSPECT. PSYCHOL. SCI. 615, at 617 (2012).

¹¹⁴ Matthew Warren, First analysis of ‘pre-registered’ studies shows sharp rise in null findings, <https://www.nature.com/articles/d41586-018-07118-1> (accessed 2019); Preregistration may also encourage researchers to think more carefully about their research design and open the door for peer scrutiny of methodological decisions: NASEM Report, *supra* note 2 at 32, 44.

¹¹⁵ Center for Open Science, *Openness of Forensic Journals*, <https://osf.io/5pk7j/> (accessed 2019).

¹¹⁶ PCAST Report, *supra* note 2 at 52.

validation researchers can make that bias their findings, such as excluding apparent outliers (eg examiners who performed very poorly) and selectively reporting the responses for certain subsets of stimuli.¹¹⁷ Moreover, the practices those are girded by this validation research impact the criminal justice system and regularly serve as inculpatory evidence in courtrooms. Effectively invisible choices that artificially lower reported error rates are immune from cross-examination and judicial gatekeeping. Preregistration would, at least, contribute to making some of these choices open to scrutiny by academics, advocates, and other stakeholders in the criminal justice process.

Given academic science's struggle with publication bias, we suspect the forensic scientific literature may also include a great many undisclosed studies that did not work out the way researchers hoped. By way of (anecdotal) example, the history of forensic bite mark identification is riddled with stories of studies conducted behind closed doors.¹¹⁸ Insider accounts are helpful in determining the results of these studies,¹¹⁹ but preregistered designs would be much more effective. Here, researchers' motivations may be problematic in both the mainstream sciences and forensic science.¹²⁰ Whereas mainstream scientists are motivated to accrue publications and citations by submitting exciting new findings (and not disclosing studies casting doubt on those findings),¹²¹ forensic scientists may be reluctant to publish results that cast doubt on their field.¹²²

Similar to preregistration, registered reports are a format of empirical article where the introduction, hypotheses, procedures, and analysis plans are submitted to a journal and peer-reviewed prior to data collection.¹²³ Peer-review of the research plan prior to data collection means that necessary revisions can be made before any resources are

¹¹⁷ Id. at 92: 'There was one false positive which the author excluded because it appeared to be a clerical error and was not repeated on subsequent retest.'; Id. at 95: 'In validation studies, it is inappropriate to exclude errors in a post hoc manner.' Similarly, preregistration of an analysis plan would help make apparent any changes made to make performance seem better. This may have been the case in a study performed by the American Board of Forensic Odontologists described by Michael Bowers in which the results were presented in a misleading way and one is difficult to conceive of as planned. See David L. Faigman et al., *Modern Scientific Evidence: The Law and Science of Expert Testimony* (2016-17) at §35:13: 'It is important to say something, first, about the meaning of the data and the way they are presented. Suppose one were told that the overall accuracy rate for a test case was 85%. One might conclude from that number that the examiners were doing reasonably well—not as well as one might hope from a forensic science that claims the ability to connect crime scene bite marks to suspects "to the exclusion of all others in the world," but not terrible either. In truth, however, the performance is far more troubling than is apparent. What is not made evident by that number is the fact that the poorest level of performance that examiners could achieve in this study—if they got every single answer as wrong as they could get it—would still make them appear to be accurate 71% of the time.'

¹¹⁸ Faigman et al, *id.* at §35:12; PCAST Report, *id.* at fn 231.

¹¹⁹ *Id.*

¹²⁰ Much can be said about incentives and motivations, and their impact on recent controversies in the mainstream and forensic sciences. For a preliminary overview, see Jason M. Chin, Bethany Grows, & David T. Mellor, *Improving Expert Evidence: The Role of Open Science and Transparency*, 50(2) OTTAWA L. REV. (forthcoming 2019), <https://osf.io/preprints/lawarxiv/t2rx6/> (last accessed 2019).

¹²¹ Indeed, Simon Cole has studied differences in the prestige economies as they exist in the mainstream and forensic sciences, see Cole, *supra* note 105.

¹²² David E. Bernstein, *Expert Witnesses, Adversarial Bias, and the (Partial) Failure of the Daubert Revolution*, 93(2) IOWA L. REV. 451 (2007); Sometimes this inclination is explicitly noted, see Mariya Goray et al., *Secondary DNA transfer of biological substances under varying test conditions*, 4(2) FORENSIC SCI. INT. GENET. 62, at 63 (2010): 'Currently there is limited knowledge concerning conditions that may influence secondary DNA transfer. This ignorance not only limits sampling strategies, DNA interpretations, and case investigations in general, it could also be easily exploited by defence councils.' [*sic*, emphasis added].

¹²³ Christopher D. Chambers, *Registered Reports: A new publishing initiative at Cortex*, 49 CORTEX 609 (2013).

expended. The article is then either rejected or receives an in-principle acceptance (ie publication is virtually guaranteed if the researchers follow the plan). One of the main benefits of registered reports is that the publication decision is based on the rigor of the methods rather than the outcome, thus curbing publication bias. Registered reports are also often used for replication research. Since the introduction of registered reports in the journal *Cortex* in 2013, a total of 126 journals spanning a wide range of scientific disciplines now accept registered reports as a publication format.¹²⁴

As with preregistrations, we did not find any forensic scientific journal that expressly mentioned registered reports or replications.¹²⁵ This is unfortunate because these reforms could be particularly useful in forensic science. A greater focus on methodology versus outcome may nudge forensic scientists towards more careful research design, creating an iterative process that improves the standards in the field.¹²⁶ Further, replication research would assist in assuring that latent experimenter effects are not biasing the existing literature.¹²⁷

Open Data, Materials, and Code

Making data, research materials, and code (eg **algorithms** performing the statistical analysis or simulation related to a study)¹²⁸ open and publicly accessible is central to the open science movement.¹²⁹ Sharing these aspects of the research process allows other researchers to confirm prior findings and detect potential error (or fabrication) in that work. Data sharing also enables researchers to combine existing data into larger datasets to perform meta-analyses and tackle novel research questions (see Part IV).¹³⁰

Despite these benefits, data has not traditionally been open. An analysis of 500 articles in 50 eminent scientific journals found that only 9% of articles had full raw data available online, despite many of the journals having policies related to open data.¹³¹ Troublingly, in 2005, when a group of researchers emailed the authors of 141 empirical articles published in the previous year to obtain raw data, 73% of the original authors were unwilling to share their data with their peers.¹³² Researchers with generally weaker results were less likely to respond to these emails.¹³³

¹²⁴ As of September 2018. For a full, regularly updated list of participating journals, see: Center for Open Science, Registered Reports: Peer review before results are known to align scientific values and practices, <https://cos.io/rr/#journals> (accessed 2019).

¹²⁵ Center for Open Science, *Openness of Forensic Journals*, <https://osf.io/Spk7j/> (accessed 2019).

¹²⁶ For example, peer review focused on methodology may force researchers to think more carefully about that aspect of the research, creating a back-and-forth that benefits the entire field.

¹²⁷ Munafò et al., *supra* note 88, at 3.

¹²⁸ Victoria C. Stodden, Trust Your Science? Open Your Data and Code, <https://web.stanford.edu/~vcs/papers/TrustYourScience-STODDEN.pdf> (accessed 2019).

¹²⁹ NASEM Report, *supra* note 2 at 28-29.

¹³⁰ Ian Hrynaskiewicz & Matthew J. Cockerill, *Open by default: a proposed copyright license and waiver agreement for open access research and data in peer-reviewed journals*, 7(5) BMC RES. NOTES. 494 (2012).

¹³¹ Alawi A. Alsheikh-Ali et al., *Public Availability of Published Research Data in High-Impact Journals*, 6(9) PLoS ONE 1 (2011).

¹³² Jelte M. Wicherts et al., *The poor availability of psychological research data for reanalysis*, 61(7) AM. PSYCHOL. 726 (2006).

¹³³ Wicherts et al., *supra* note 58.

Like with preregistration, journals can promote open data, materials, and code. As we noted above, several TOP standards cover these aspects of the research process.¹³⁴ By way of example, a TOP signatory journal, *Science*, recently updated its editorial policy to require authors to make their data available, subject to ‘truly exceptional circumstances’.¹³⁵ Attitudes among researchers seem to be tracking these updated editorial policies. The 2017 *State of Open Data Report* found that awareness of open data sets and researchers’ willingness to use open datasets were positively trending.¹³⁶

Increases in open data may also be due to better infrastructure. For example, the OSF allows researchers to upload materials, datasets, and code organized under the same project with a persistent Digital Object Identifier (DOI). Other popular cross-disciplinary open data repositories include Figshare, Zenodo, and the Harvard Science Framework.¹³⁷ Further, Google recently launched a new initiative, Dataset Search, to help researchers find open data. This works similarly to Google Scholar as it accesses datasets from publisher’s websites, personal websites, and institutional repositories.¹³⁸

We were encouraged to see that, unlike with preregistration, forensic scientific journals appear somewhat concerned with transparency of data and code (see Table 1). Our findings show that 15 of 30 journals encouraged data transparency and 11 encouraged code transparency. Still, they remain at TOP level 0 (ie mere encouragement) on this standard (and have not formally adopted TOP). As with preregistration, we believe that opening the research process will benefit forensic science in the long run: sharing of data and materials provides much efficiency and promotes error correction. Furthermore, from a criminal justice perspective, we would question the fairness of asking the criminally accused to simply trust the closed forensic scientific literature knowing what has occurred in the mainstream sciences. Still, openness itself may present significant legal issues (eg privacy).¹³⁹

Open Access Journals

Finally, open access to journal articles has been a contentious issue for decades, inspiring some of the first discussions about open science.¹⁴⁰ Typically, published articles

¹³⁴ They are standards (ii) through (v): data transparency, analytic methods (code) transparency, research materials transparency, and design and analysis transparency.

¹³⁵ Marcia McNutt, Taking up TOP, 352(6290) *Science* 1147 (2016); Science, Science Journals: Editorial Policies, <https://www.sciencemag.org/authors/sciencejournals-editorial-policies> (accessed 2019) [Science Editorial Policy]; The Royal Society, Data sharing and mining, <https://royalsociety.org/journals/ethics-policies/data-sharing-mining/> (accessed 2019).

¹³⁶ Figshare, The State of Open Data Report 2017, https://figshare.com/articles/The_State_of_Open_Data_Report_2017/5481187/1 (accessed 1 October 2018). More foundationally, the NASEM Report points to the FAIR Guiding Principles as a model for data openness. Developed by an international body of academics, industry partners, funders, and publishers, the FAIR guidelines seek to ensure that research products such as data, materials, and code are findable, accessible, interoperable, and reusable. See NASEM Report, *supra* note 2 at 28–29; Global Open FAIR, GO FAIR: a bottom-up international approach, <https://www.go-fair.org> (accessed Sep. 22, 2018).

¹³⁷ Figshare, <https://figshare.com> (accessed 2019); Harvard Dataverse, <https://dataverse.harvard.edu> (accessed 2019); Zenodo, <https://zenodo.org> (accessed 2019).

¹³⁸ Natasha Noy, *Making it easier to discover datasets*, Google Blog (9 September 2018), <https://www.blog.google/products/search/making-it-easier-discover-datasets/>.

¹³⁹ NASEM Report, *supra* note 2 at 50–52. We discuss these concerns in Part IV.

¹⁴⁰ Budapest Open Access Initiative, <https://www.budapestopenaccessinitiative.org/read> (accessed Sep. 8, 2018); John Willinsky, *Scholarly Associations and the Economic Viability of Open Access Publishing*, 4(2) *JoDI* 1 (2004). See *Paywall* documentary, *supra* note 53.

are only available to those with (costly) subscriptions. However, there is now a trend towards making articles open access, either through fully open access journals¹⁴¹ or hybrid journals which charge authors a fee to make their article open access, if they wish.¹⁴² There is much variation in open access among disciplines, with the life and biomedical sciences embracing open access and several fields such as the social sciences and professional fields lagging behind.¹⁴³ In addition to allowing greater public access to science, research has demonstrated that articles in open access journals are more likely to be downloaded and cited.¹⁴⁴ Free servers also exist to allow researchers to post preprints of their research (eg LawArXiv in law and PsyArXiv in psychology).¹⁴⁵ We are not aware of a preprint service dedicated to forensic science.

Forensic scientific journals generally provide open access options to authors (see Table 1). We only found three journals with no open access option at all. One new journal with an open focus provides *only* the option of open access publishing.¹⁴⁶ Open access publishing in forensic science is incredibly important. Many stakeholders in the criminal justice system cannot be expected to have access to academic subscriptions (and this likely explains why so many forensic journals have open access options). This includes defense lawyers, accused parties, and forensic scientists themselves (who often are not affiliated with a university). An important issue going forward will be keeping author publishing charges manageable, especially given the limited grant funding available to forensic science researchers.¹⁴⁷

PART IV. OPEN FORENSIC SCIENCE

Open scientific reform offers several distinctive advantages to forensic science, a field that endeavors to see justice done while avoiding error. In this section, we will survey three general ways in which openness can improve forensic science: establishing the validity of existing methods, developing new objective methods, and applying those methods in a trustworthy way (see Table 2). We will end with a discussion of the barriers these reforms will face.

¹⁴¹ For a full list of open access journals, see Directory of Open Access Journals, <https://doaj.org> (accessed 2019). 'Green' open access articles are available in a public repository, possibly after some embargo period. 'Gold' open access articles are available freely from the publisher.

¹⁴² Julia Frankland & Margaret A. Ray, *Traditional versus Open Access Scholarly Journal Publishing: An Economic Perspective*, 49(1) J. SCH. PUBL. 5 (2017).

¹⁴³ Jeroen Bosman & Bianca Kramer, Open access levels: a quantitative exploration using Web of Science and oaDOI data, PeerJ Preprint (8 September 2018), <https://peerj.com/preprints/3520/>; NASEM Report, *supra* note 2 at 59-63.

¹⁴⁴ Philip M. Davis et al., *Open access publishing, article downloads, and citations: randomized controlled trial*, 337(1) BMJ 1 (2008); Philip M. Davis, *Open access, readership, citations: a randomized controlled trial of scientific journal publishing*, 25(7) FASEB J. 2129 (2011); A. Ben Wagner, *Open Access Citation Advantage: An Annotated Bibliography*, 60(1) ISSUES SCI. TECHNOL. Librariansh. 1 (2014).

¹⁴⁵ LawArXiv, <http://lawarxiv.info/>; PsyArXiv, <https://psyarxiv.com/>; Munafò et al., *supra* note 88, at 6; NASEM Report, *supra* note 2 at 69-74.

¹⁴⁶ Elsevier, Forensic Science International: Synergy – Open Access Journal, <https://www.elsevier.com/journals/forensic-science-international-synergy/2589-871X/open-access-journal> (last accessed 2019).

¹⁴⁷ NAS Report, *supra* note 2 at 187.

Table 2. A list of the initiatives we have recommended, the goal within forensic science they accord with, and the benefits that will accrue if they are adopted.

<i>Forensic science goal</i>	<i>Recommended initiatives</i>	<i>Benefits</i>
Foundational validity	Preregistration	Controlling questionable research practices (QRPs), reducing experimenter bias, reducing false positive results
	Registered reports	Controlling QRPs, reducing experimenter bias, reducing publication bias, reducing false positive results
	Replication	Reducing false positive results, reducing publication bias, reducing experimenter bias
	Multi-center collaborative studies (eg Many Labs)	Promoting collaboration, isolating setting and experimenter effects, reducing type M errors
Objective methods	Establishing ForensicsArXiv server	Reducing publication bias, faster dissemination of results, research available to legal actors and forensic practitioners
	Large, open source databases	Promoting collaboration, large ground truth stimuli set, ability to test examiners and algorithms using ground truth stimuli
Applied validity	Preregistering analytic choices	Controlling and revealing unconscious biases, accountability
	Open workflow and analysis	Controlling and revealing unconscious biases, accountability
	Open proficiency testing and error repositories	Accurate error measurements, accountability

Establishing Foundational Validity Through ‘Many Labs’

An immediate and fundamental challenge facing forensic science is establishing the validity of many of its methodologies.¹⁴⁸ For subjective methods, which many forensic practices still are, the PCAST Report recommended large-scale ‘black-box’ studies of performance in situations in which the ground truth is known.¹⁴⁹ In other words, we cannot know what is going on in the black-box of the examiner’s brain. We can, however, infer that those subjective processes are working as expected if we expose *many examiners* to many samples that come from known sources, and measure how often they come to the correct answer. As we discussed above, this type of research has been surprisingly uncommon, in part, because it is resource-demanding.

An amalgam of preregistration, registered reports, and replication—increasingly used in psychological research—may provide a paradigm for forensic science to follow in its validation efforts. Psychology, a relatively early-embracer of open science reforms,¹⁵⁰ shares many of forensic science’s struggles. Like the measurement of subjective forensic expertise, psychology often seeks to measure qualia. This poses many challenges, including the fact that individuals, unlike chemicals and atoms, vary in difficult-to-predict ways.¹⁵¹ False positive and negative results may therefore result from sampling variation and measurement error.

To overcome the inherent challenges in measuring subjective processes in psychology, some researchers are relying on multi-center collaborative studies that have historically been used in some medical and genetic association research.¹⁵² One successful model is the ‘Many Labs’ replication projects (see also the Pipeline Project, the Psychological Science Accelerator, the Collaborative Replications and Education Project, and Study Swap).¹⁵³ In these studies, the project leads begin by identifying a controversial or highly cited finding and seeking collaborators on the OSF or through their existing networks. The group may then consult with stakeholders like the party that initially discovered the contested finding and eventually agree on and preregister a protocol. The individual labs then recruit participants and run the protocol, each producing results that can be both pooled between labs and analysed individually or by a third party.¹⁵⁴

Many Labs style projects offer a host of benefits. As we discussed above, preregistration is important in controlling QRPs and publication bias. Replication across labs also helps to isolate effects related to the setting of the study (eg whether examiners

¹⁴⁸ PCAST Report, *supra* note 2.

¹⁴⁹ *Id.* at 50–54.

¹⁵⁰ NASEM Report, *supra* note 2 at 44: ‘It may take time for research communities to transition to open practices that enable wider review and scrutiny of research. Psychology is a current encouraging example.’

¹⁵¹ But see Baker, *supra* note 56.

¹⁵² Munafò et al., *supra* note 88, at 3.

¹⁵³ For Many Labs, see Many Labs 1 & 3, *supra* note 60; For the Pipeline Project, see Martin Schweinsberg, *The pipeline project: Pre-publication independent replications of a single laboratory’s research pipeline*, 66 J. EXP. SOC. PSYCHOL. 55 (2016); For the Psychological Science Accelerator, see Hannah Moshontz et al., *The Psychological Science Accelerator: Advancing Psychology Through a Distributed Collaborative Network*, AMPPS (2018); For the Collaborative Replications and Education Project see <https://www.psychologicalscience.org/observer/replication-education>; For Study Swap, see Center for Open Science, *StudySwap: A platform for interlab replication, collaboration, and research resource exchange*, <https://osf.io/view/StudySwap/> (accessed 2019); Jon Grahe et al., *Replication Education*, Association for Psychological Science (APS) (accessed 2019).

¹⁵⁴ One challenge in determining the false positive rate of fingerprint analysis is that experimenters have all used slightly different protocols, see PCAST Report, *supra* note 2 at 95.

trained in a particular lab outperformed others) and any latent experimenter effects.¹⁵⁵ Importantly, the large sample sizes provided by Many Labs projects contribute to control of ‘Type M’ errors, or errors related to estimating the *magnitude* of a study’s effect. As influential statisticians have noted, the mainstream sciences have regularly been concerned with false positives and negatives, often overlooking Type M errors.¹⁵⁶ Recall the large-scale 2018 effort to reproduce the outcomes of 21 studies published in *Nature* and *Science* mentioned above. The researchers in that study found that effect sizes were 50% smaller than in the original studies—considerable Type M error.¹⁵⁷

Type M errors are especially important in the foundational forensic literature because courts require *precise* estimates of a method’s error rate to ascertain its probative value. For example, research is converging to demonstrate that expert fingerprint examiners considerably outperform laypeople in identifying the source of a fingerprint.¹⁵⁸ There is still, however, considerable variance in the estimates of their error (eg one false positive in 24 judgments to one in 604).¹⁵⁹ Factfinders ought to be provided with accurate estimates, which large collaborative projects can help provide. Moreover, as we have noted, independent replication is central to the scientific process. Despite this principle, the PCAST Report declared fingerprint analysis to be foundationally valid on the basis of only two studies (both performed by law enforcement agencies).¹⁶⁰ Adopting a Many Labs approach in forensic science may lend confidence to the Report’s conclusion.

In projects like Many Labs, open science reformists note the importance of independent methodological support. Such a mechanism may be especially useful in forensic science, in which the quality of methodological training has often been unevenly distributed among practitioners, researchers, and those in hybrid roles.¹⁶¹ Here, NIST may play a role similar to the successful experience found in the case of the establishment of the Independent Statistical Standing Committee (ISSC) by researchers of Huntington’s disease.¹⁶² Members of the ISSC have strong methodological training and, importantly, no interest in the outcome of research into the disease’s treatment. The Committee’s role has been expanded since its establishment.

Transforming Subjective Into Objective Methods

Beyond validating subjective methods, forensic science is also moving towards developing and validating objective methods. Great strides, for instance, have been made

¹⁵⁵ For instance, the PCAST Report, *id.* at 79-82 called for independent replication of complex mixture DNA analysis by parties ‘not associated with the software developers’ [emph. in original]. It also, at 95, noted the difficulty of comparing some fingerprint studies because they used different methodologies and materials. Coordination in the fashion of the Many Labs studies could help control these factors.

¹⁵⁶ Andrew Gelman & John Carlin, *Beyond Power Calculations: Assessing Type S (Sign) and Type M (Magnitude) Errors*, 9(6) PERSPECT. PSYCHOL. SCI. 641 (2014).

¹⁵⁷ Camerer et al., *supra* note 61.

¹⁵⁸ Jason M. Tangen et al., *Identifying Fingerprint Expertise*, 22(8) PSYCH. SCI. 995 (2011).

¹⁵⁹ See PCAST Report, *supra* note 2 at 98.

¹⁶⁰ *Id.* at 101.

¹⁶¹ NAS Report, *supra* note 2 at 218-239.

¹⁶² See Munafò et al., *supra* note 88, at 4. Further, the PCAST Report, *supra* note 2 at 17 suggested that the FBI Laboratory expand its collaboration with external scientists.

using automated image analysis to perform fingerprint identification.¹⁶³ Going forward, the most important resource this initiative needs is access to ‘huge databases containing known prints’.¹⁶⁴ Similarly, the development of objective methods to associate ammunition with a specific firearm (ie toolmark analysis) and analysis of complex DNA mixtures is similarly hampered by lack of a sufficiently large database.¹⁶⁵ The PCAST Report lamented the fact that the FBI has not opened many of its databases, including those with no privacy concerns (eg toolmarks).¹⁶⁶

Despite some hesitation, some programs—founded in open scientific principles—are already underway to develop objective methods using open databases. For example, the PROVEDIt Initiative has made available 25,000 DNA profiles from mixed sources that can be used to validate DNA analysis software.¹⁶⁷ Similarly, an industry partnership between the University of New South Wales and the Australian Passport Office is crowdsourcing ground truth facial images to test the accuracy of facial recognition algorithms through the #Selfies4Science program.¹⁶⁸ These are all promising developments, but they could be augmented by grassroots sharing of materials by individual laboratories through systems like the OSF. The collaboration behind #Selfies4Science, for instance, has not made their database available to other researchers.

Improving Applied Validity Through Openness and Transparency

Most of the reforms we have discussed so far involve *conducting* research more transparently. In forensic science, however, there is also the matter of putting that research into practice. As the PCAST Report said, forensic scientific disciplines must be both foundationally valid and applied in a valid way. As to applied validity, open science reforms and initiatives are less directly applicable. But still, some open principles and techniques can be applied to forensic scientific practice.¹⁶⁹ We will discuss three: (1) transparently reporting forensic analytic choices; (2) open forensic workflow and analysis; and, (3) open proficiency testing and error repositories. Central to all of our suggestions is transparency and removing some discretion in what practitioners report about their process. As we have seen, even well-trained academic scientists have used flexibility in their methods to generate misleading results. We should be concerned about the same issues occurring in applied forensic scientific practice.

First, consider employing greater transparency in forensic practice, particularly fingerprint analysis. As part of their methodology, fingerprint examiners determine which features or ‘minutiae’ of a latent print (ie one found at a crime scene) are distinctive and

¹⁶³ This tool assists with fingerprint identification by providing quantitative values for the likelihood of identifying the same features in another print (like random match probabilities in DNA). It still relies on the analyst to identify features for comparison and make the comparison conclusion. PCAST Report, *supra* note 2 at 10–11, 103; Henry J. Swofford et al., *A method for the statistical interpretation of friction ridge skin impression evidence: Method development and validation*, 287(1) FORENSIC SCI. INT. ¹¹³ (2018).

¹⁶⁴ PCAST Report, *id.* at 11.

¹⁶⁵ PCAST Report, *id.* at 82, 114.

¹⁶⁶ *Id.* at 132–133.

¹⁶⁷ Alsonse et al., *A large-scale dataset of single and mixed-source short tandem repeat profiles to inform human identification strategies: PROVEDIt*, 32 Forensic Sci. Int. Genet. 62 (2017); Boston University, *Access and download .fsa, .hid or .csv PROVEDIt Database Files*, <http://www.bu.edu/dnamixtures/pages/help/downloads/> (accessed 2019).

¹⁶⁸ See UNSW Sydney, #Selfies4Science, <https://www.selfies4science.com> (accessed 2019).

¹⁶⁹ For a review of the transparency of state crime labs, see Cásarez & Thompson, *supra* note 2.

will thus be important during comparison.¹⁷⁰ However, practitioners—*after* viewing the comparison (ie exemplar) print—can go back and alter those features they deemed important. This practice, if not fully documented—and it often is not—can be highly misleading and result in undisclosed confirmation bias and circular reasoning.¹⁷¹

We suggest that examiners be required to transparently document the features they predict will be diagnostic during the analysis stage.¹⁷² Langenburg and Champod have developed a color-coded system—the Green-Yellow-Red-Orange (GYRO) system—for fingerprint examiners to document their analytic choices.¹⁷³ If an examiner is highly confident in the existence of that feature in the latent print and has a high expectation that the feature will be present in an exemplar print, the examiner will mark that feature with green. If the examiner has a medium or low level of confidence in that feature, they will mark it with yellow or red, respectively. Finally, the orange color represents features that were not identified during the initial analysis of the latent print, but were observed after viewing the exemplar print.

Here we do not mean to constrain examiners. Indeed, there may be cases in which such re-analysis is beneficial: an examiner may have incorrectly discounted a genuine feature as being an artefact but, upon seeing the same feature in the exemplar print, may realize that it was indeed diagnostic.¹⁷⁴ However, much like a strikethrough on incorrect case note documentation, the examiner should also document edits to his or her feature analysis. As in the mainstream sciences, it is important to be open and candid about the reality of the process and the serious opportunities for bias to creep into it.

More transparent and open analytic choices flows into our second and more general point: forensic laboratories should operate on the principle of transparency.¹⁷⁵ Several aspects of forensic practice involve more discretion and subjectivity than judges and jurors would reasonably expect, and more than is admitted in expert reports.¹⁷⁶ For

¹⁷⁰ PCAST Report, *supra* note 2 at 10, 91; Alicia Rairden et al., *Resolving latent conflict: What happens when latent print examiners enter the cage?*, 289 FORENSIC SCI. INT. 215, at 216 (2018); Bradford T. Ulery et al., *Changes in latent fingerprint examiners' markup between analysis and comparison*, 247(1) FORENSIC SCI. INT. 54 (2014). See the following document for a more comprehensive explanation of each stage of the ACE-V procedure: Expert Working Group on Human Factors in Latent Print Analysis, *Latent Print Examination and Human Factors: Improving the Practice through a Systems Approach* (2012).

¹⁷¹ PCAST Report, *supra* note 2 at 10. We stress that there may be valid reasons for revisiting earlier portions of a fingerprint analysis based on information gleaned later in the analysis. However, this process should be fully documented.

¹⁷² We have used latent fingerprint analysis to exemplify this point, however the same can be done for many other comparative forensic sciences. For example, unknown or questioned DNA samples should be interpreted and documented before comparison with a reference sample.

¹⁷³ G. Langenburg & Christophe Champod, *The GYRO System — A Recommended Approach to More Transparent Documentation*, 61(4) J. FORENSIC IDENT. 373 (2011).

¹⁷⁴ Expert Working Group on Human Factors in Latent Print Analysis, *Latent Print Examination and Human Factors: Improving the Practice through a Systems Approach*, at 42 (2012).

¹⁷⁵ Although the state of transparency among crime labs is generally poor, there may be change on the horizon. In 2018, Deputy Attorney General Rod Rosenstein committed to making transparency a 'core value': Rod J. Rosenstein, Deputy Attorney General Rosenstein Delivers Remarks at the American Academy of Forensic Sciences (Feb. 21, 2018) <https://www.justice.gov/opa/speech/deputy-attorney-general-rostenstein-delivers-remarks-american-academy-forensic-sciences> (last accessed 2019). See also Cásarez & Thompson, *supra* note 2.

¹⁷⁶ See Mnookin, *supra* note 2, at 1226; Gary Edmond, David Hamer, & Emma Cunliffe, *A little ignorance is a dangerous thing: engaging with exogenous knowledge not adduced by the parties*, 25(3) GRIFFITH L. REV. 383 (2016).

instance, there is considerable variation in practices between labs about whether examiners should verify the work of other examiners blind to the original decision, whether the first examiner can choose the verifying examiner, and if discussions are permitted or encouraged between these individuals.¹⁷⁷

Additionally, thorough documentation should be conducted during verification and/or technical review procedures wherein analytical conclusions come into question. Discussions between analysts inevitably influences one or both analyst much like the aforementioned fingerprint example of editing initial results once a comparison exemplar is provided. Without proper and thorough documentation of consultation between forensic practitioners, the true nature of a peer reviewed result may not be apparent or available for future analysis. While some forensics labs are adopting very transparent protocols surrounding these decisions, such openness has not yet reached orthodoxy (or even come close to this).¹⁷⁸ Indeed, in our (anecdotal) conversations with forensic examiners, several have expressed hesitation at disclosing aspects of their analysis that could convey a lack of certainty (eg that analyzing a certain latent print took longer than others). This also leads us to question their openness when they are questioned in court about verification, consultation and conflict procedures.

To remedy these problems, we suggest that laboratories freely publish their standard operating procedures, as well as any analytical methodology utilized.¹⁷⁹ Such steps have been taken by both the Houston Forensic Science Center (HFSC) and the Idaho State Police (ISP) who publish these data on their public websites.¹⁸⁰ The HFSC and ISP both have public facing websites where the public can review analytical methods and accreditation information (eg the ISP posts staff CVs and the HFSC has plans to do so in the future). The HFSC also publishes standard operating procedures, calibration, instrumentation records, batch records, quality reports, and incident/preventative action/corrective action reports. Specifically, these data could be utilized by another party to re-analyze the evidence to see if the results are the same (however, they likely would have to use the same instrument for exact findings because some analytical results can be impacted by instrumentation). Moreover, sharing details of crime lab operating procedures may result in efficiencies as labs can learn from the successes and challenges of other labs.¹⁸¹ Pursuant with open science standards, the HFSC and ISP will eventually have to consider who will be the long-term stewards of this data.¹⁸²

¹⁷⁷ Kaye N. Ballantyne et al., *Peer review in forensic science*, 277 FORENSIC SCI. INT. 66, at 70-72 (2017): "The discussions between peer reviewers, changes made to opinions and statements as a result of peer review are, to our knowledge, rarely disclosed in reports and oral testimony, although they may be included in the case file."

¹⁷⁸ For a positive example, see the self-scrutiny apparent in the practices of the Houston Forensic Science Center, <http://www.houstonforensicscience.org/index.php> (accessed 2019), Rairden et al., *supra* note 173, or their website. For the orthodox position, see Saul M. Kassir et al., *The forensic confirmation bias: Problems, perspectives, and proposed solutions*, 2(1) J. APPL. RES. MEM. COGN. 42 (2013); Rachel A. Searston et al., *Putting bias into context: The role of familiarity in identification*, 40(1) LAW. HUM. BEHAV. 50 (2016); Nikola K.P. Osborne & Rachel Zajac, *An imperfect match? Crime-related context influences fingerprint decisions*, 30(1) APPL. COGN. PSYCHOL. 126 (2016).

¹⁷⁹ See Cásarez & Thompson, *supra* note 2, at 1014-1030.

¹⁸⁰ See Houston Forensic Science Center, Houston Forensic Science Center RecordSearch, <https://records.hfscdiscovery.org> (accessed Dec. 7, 2018); Idaho State Police, Accreditation & Staff CV's, <https://www.isp.idaho.gov/forensics/pillsPages/resumes.html> (accessed Dec. 7, 2018).

¹⁸¹ Cásarez & Thompson, *supra* note 2.

¹⁸² NASEM Report, *supra* note 2 at 10.

Third and finally, more objective measures of error do exist and are useful in ascertaining the probative value of forensic evidence. More attention should be paid towards making these measures both as open and as useful as they can be (goals that often converge). For example, forensic examiners take proficiency tests, which determine how well they can apply a particular technique.¹⁸³ However, these proficiency tests are typically commercially obtained and thus not open to scrutiny from the broader scientific community. Moreover, they typically do not mimic routine casework and are therefore non-blind. Nonblind proficiency tests are problematic as forensic analysts may not behave in the same manner as for routine casework due to the knowledge of being tested, thereby skewing examiner error rate.¹⁸⁴ The PCAST Report strongly recommended using proficiency testing, but urged testing services to publicly release the tests so that other scientists could determine if the testing is realistic.¹⁸⁵

Beyond the individual practitioner, labwise error rates inform the value of forensic expertise and should be provided more openly. Eschewing a culture that denies the possibility of error,¹⁸⁶ some labs are beginning to implement policies of radical transparency by publicly reporting mistakes (but labs appears more reluctant to report errors than to adopt other transparency-related reforms).¹⁸⁷ The beginnings of change here in forensic science are analogous to previous practices in the mainstream sciences. The previous closed model of science would actively suppress studies that did not fit the experimenter's narrative.¹⁸⁸ Now they are being reported in a move that, in our view, improves the credibility of science.

Barriers to Open Forensic Science

While we have struck an optimistic tone in our analysis of open science's applicability to forensics, there are certainly substantial barriers to any vision of open forensic science (just as there are with open science generally).¹⁸⁹ We believe the advantages of open science make addressing these barriers worthwhile. Still, the challenges in implementing many of the reforms we have described deserve careful consideration.

One possible resistance point to embracing openness is the culture of forensic science, which tends to resist admitting errors.¹⁹⁰ It will therefore be challenging to

¹⁸³ PCAST Report, *supra* note 2 at 57-58; Collaborative Testing Services, Inc., Collaborative Testing Services is your Proficiency Testing Expert, <https://cts-forensics.com/index-forensics-testing.php> (accessed Oct 12, 2018); However, see Jonathan J. Koehler, *Forensics or Fauxrensic? Ascertaining Accuracy in the Forensic Sciences*, 49 ARIZ. ST. L.J. 1369, at 1395 (2018) for a review of why current proficiency testing efforts fall short: 'CTS tests do not use realistic samples, do not use blind testing, and they do not control the way laboratories or examiners use their tests. Nevertheless, courts continue to rely heavily on these data ... to justify the conclusion that error rates are sufficiently low.' Additionally, the PCAST Report at 68 notes that 'the forensic community disfavors more challenging tests—and that testing companies are concerned that they could lose business if their tests are viewed as too challenging'.

¹⁸⁴ Cásarez & Thompson, *supra* note 2, at 1063-1065.

¹⁸⁵ PCAST Report, *supra* note 2 at 58.

¹⁸⁶ Ballantyne et al., *supra* note 177, at 72.

¹⁸⁷ Cásarez & Thompson, *supra* note 2, at 1027-1030; PCAST Report, *supra* note 2 at 18, 74.

¹⁸⁸ Marjan Bakker et al., *The Rules of the Game Called Psychological Science*, 7(6) PERSPECT. PSYCHOL. SCI. 543 (2012).

¹⁸⁹ On the barriers to openness in the mainstream sciences, see the NASEM Report, *supra* note 2 at 37-58.

¹⁹⁰ See Simon Cole, *More than Zero: Accounting for Error in Latent Fingerprint Identification*, 95(3) J. CRIM. L. & CRIMINOLOGY 985 (2005).

promote transparency about mistakes that do inevitably happen and research that does not fit with the experimenter's hypothesis. One way to advance this aim may be through increased partnerships between academic scientific labs that have embraced open science reforms and forensic scientists. A step in this direction can be seen in the American Academy of Forensic Sciences Laboratories and Educators Alliance Program (LEAP), which aims to connect academia and forensic laboratories.¹⁹¹ Federal organizations in the US may also wish to fund similar joint projects. This may not just produce strong research, but also contribute to training and education in open research methods for forensic professionals.

Secondly, transitioning to open research involves significant financial costs, at least in the beginning. For example, the NASEM Report anticipated challenges in shifting the current publishing system away from a subscription-based model.¹⁹² Most notably, and as we have seen in the forensic scientific journals (see the discussion of open access in Part III), author publishing fees can be substantial and possibly prohibitive for many researchers.

Going forward, the NASEM Report anticipated that reasonable publishing costs may be incorporated into grant funding.¹⁹³ Indeed, funders are beginning to acknowledge the importance of open scholarly communication and even require that applicants plan to make their work freely available.¹⁹⁴ Agencies funding foundational forensic science research ought to be especially attuned to making the fruits of that labor open. Unlike academic scientists who typically have access to journal articles through their institution, forensic practitioners and lawyers often do not have this luxury. As we noted above, it is incredibly important that practitioners have access to studies providing foundational research and new insights. However, as technology improves and competition increases, we may also expect publishing fees to decrease. Interstitially, forensic scientists may wish to publish their work as preprints, perhaps through the development of a Forensic ArXiv server.

Beyond publishing, forensic scientific researchers may find economies by leveraging platforms and programs already developed in the open science movement. As we have discussed, the OSF and more specific initiatives like Many Labs provide useful infrastructures for forensics to build on.

From an economic standpoint, there is also an issue with companies claiming trade secret privilege over the workings of forensic scientific software.¹⁹⁵ This issue is exacerbated when such technologies rest on machine learning algorithms that become a black-box because they have evolved beyond their original programming. Such software should be carefully validated. And if designers are not willing to disclose the

¹⁹¹ American Society of Crime Laboratory Directors, *Forensic Research Committee*, <https://www.asclld.org/forensic-research-committee/> (accessed 2019). See also our discussion of a recent Australian Research Council-funded collaboration between a university psychology laboratory and several police departments, which is using open science methods to collaborate and conduct research: UQ News, *supra* note 41.

¹⁹² NASEM Report, *supra* note 2 at 37-43.

¹⁹³ *Id.* at 136.

¹⁹⁴ See the description of the Holdren Memo in NASEM Report, *id.* at 128.

¹⁹⁵ PCAST Report, *supra* note 2 at 78-80; Rebecca Wexler, *Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System*, 70(S) STAN. L. REV. 1343 (2018).

program in sufficient detail, courts may wish to limit the admission of the results of such tests.¹⁹⁶

Finally, perhaps the most challenging issues facing open forensic science are those concerning privacy and security.¹⁹⁷ In the open science movement, these issues have provoked a great deal of discussion. For instance, the NASEM Report acknowledged that the interests of patient confidentiality and national security may provide good cause to limit the scope of open science in some cases.¹⁹⁸

When it comes to opening forensic science, the exigency of privacy and security depends on the practice and context being considered. For example, sharing materials between research groups that are not associated with individuals (eg toolmarks) do not evoke obvious privacy concerns.¹⁹⁹ They may, however, have security consequences by providing adversaries insight into investigative techniques. As to the labs applying such research and providing the public more transparency about their processes, they will have to think carefully about when to limit that information (and in some cases this is already occurring).²⁰⁰

On the other end of the spectrum to practices like toolmark analysis are practices that aim to identify individuals (eg fingerprints, DNA). These practices are not themselves uniform in the privacy and consent issues they raise.²⁰¹ For instance, DNA diverges from fingerprints as, beyond providing identifying information, it also carries a great deal of personal genetic information about the individual and his or her family. Indeed, recent advances in mapping the human genome have resulted in considerable debate about protecting genetic information (ie genetic privacy).²⁰² Note, however, that unlike more controversial research, current forensic DNA analysis practices do not rely on whole-genome sequencing.²⁰³ In fact, the field's current knowledge of DNA analysis and existing validation studies have been greatly aided by some level of open science, both through access to government databases and collaborations between researchers providing samples from local populations (and recall our above discussion of the PROVEDIt database for validating mixed source DNA analysis).²⁰⁴ Still, as

¹⁹⁶ This occurred in *State v. Pfenning*, No. 57-4-96, slip op. at 52-54, 68 (Vt. Dist. Ct. Apr. 6, 2000). See generally Edward K. Cheng & Alex Nunn, *Beyond the Witness: Bringing A Process Perspective to Modern Evidence Law*, 97(6) TEXAS L. REV. (forthcoming 2019).

¹⁹⁷ For a review of such challenges in open science generally, see NASEM Report, *supra* note 2 at 50-53.

¹⁹⁸ NASEM Report, *id.* at 2: 'Sharing data, code, and other research products is becoming more common, but barriers related to ensuring patient confidentiality and the protection of national security information exist in some domains. Proprietary research also presents barriers. Ultimately, some parts of the research enterprise may not be open'.

¹⁹⁹ PCAST Report, *supra* note 2 at 11-12.

²⁰⁰ See the practices in place at the Houston Forensic Science Center, Cásarez & Thompson, *supra* note 2.

²⁰¹ D.H. Kaye, *Behavioral Genetics Research and Criminal DNA Databases: Laws and Policies*, in THE IMPACT OF BEHAVIOURAL SCIENCES CRIMINAL LAW 362 (Nita A. Farahany ed., 2009): 'It is the samples, not the essentially random numbers contained in the databases, that pose a serious privacy question and that make DNA database systems more threatening than, say, fingerprint databases'.

²⁰² See Jane Kaye, *The Tension Between Data Sharing and Protection of Privacy in Genomics Research*, 13(1) ANNU. REV. GENOMICS & HUM. GENET. 415 (2012); Yaniv Erlich & Arvind Narayanan, *Routes for breaching and protecting genetic privacy*, 15 NAT. REV. GENET. 409 (2014).

²⁰³ Whole-genome sequencing, due to the depth of information it carries, is especially fraught with privacy issues, see: Kaye, *Id.*

²⁰⁴ For research comparing samples to government databases see Jo-Ann Bright et al, *Searching mixed DNA profiles directly against profile databases*, 9 FORENSIC SCI. INT. GENET. 102 (2014). For collaborative DNA analysis

technology improves, it will always be possible that identifying information will be (mis)used in ways that cannot be currently foreseen.²⁰⁵

Despite the risks, potential threats to privacy and security should not simply end the conversation about opening some forensic science practices. Rather, it should inspire thoughtful legal-scientific policy research seeking to progress science, while respecting privacy. In the case of open forensic science, some insights may come through conversations and collaborations between those wrestling with these issues in the open science domain.²⁰⁶ Useful models may be found in the thorough consent framework used by the Personal Genome Project and the Precision Medicine Initiative, in which volunteers share their genomic data and personal health data, respectively.²⁰⁷ It should be noted that these models are still in their infancy and remain controversial.²⁰⁸

PART V. CONCLUSION

Science—our culture's principal means of answering factual questions—is changing. It is being conducted more openly and transparently. There are many reasons for these reforms: open science is more democratic and inclusive; it enables more thorough assessment of factual claims; and, it facilitates more collaborative and efficient research. The direct impetus for many of the reforms going on in science was a crisis of confidence: opaquely conducted science was producing results that could not be reproduced.

A similar crisis of confidence may be engulfing forensic science. Attentive researchers have long noted the surprising frequency at which forensic science has committed factual mistakes. Media attention and subsequent popular knowledge seems to be catching up with this academic research.²⁰⁹ When law—a field inextricably tied to

research, see Bruce Budlowle et al, *CODIS STR Loci Data from 41 Sample Populations*, 46(3) J. FORENSIC SCI. 453 (2001). In the United States, the FBI's CODIS (Combined DNA Index System) coordinates state and federal DNA databases that can be used for investigatory purposes. As for research, the DNA Identification Act provides that the information in CODIS must be maintained such that stored DNA samples and analyses are only disclosed for identification purposes, in judicial proceedings, for criminal defense purposes, and 'if personally identifiable information is removed, for a population statistics database, for identification research and protocol development purposes, or for quality control purposes'. DNA Identification Act 42 U.S.C. § 14132(b)(3)(D) (2000).

²⁰⁵ One example may be issues that have arisen around familial matching (ie seeking inexact matches between a found DNA profile and profiles on an offender DNA database in the hopes of finding a close relative of the offender – and then using that relationship to identify the offender). For a review, see Joyce Kim et al, *Policy implications for familial searching*, 2(1) INVEST. GENET. 22 (2011).

²⁰⁶ NASEM Report, *supra* note 2 at 50-53.

²⁰⁷ Misha Angrist, *Eyes wide open: the personal genome project, citizen science and veracity in informed consent*, 6(6) PERS. MED. 691 (2009); NASEM Report, *id.* at 93-94; The White House, *Precision Medicine Initiative: Privacy and Trust Principles*, <https://obamawhitehouse.archives.gov/sites/default/files/microsites/finalpmiprivacyandtrustprinciples.pdf> (accessed 2019) (cited in NASEM Report at 94); Jeantine E. Lunshof et al, *From genetic privacy to open consent*, 9(5) NAT. REV. GENET. 406 (2008).

²⁰⁸ Zubin Master et al., *Biobanks, consent and claims of consensus*, 9(9) NAT. METHODS 885 (2012). See also Kaye, *supra* note 205; Timothy Caufield & Blake Murdoch, *Genes, cells and biobanks, Yes there's still a consent problem*, 15(7) PLOS BIOL. (2017); Clarissa Allen, Yann Joly, & Palmira Granados Moreno, *Data sharing, biobanks and informed consent: A research paradox?*, 7(1) MCGILL J. L. & HEALTH 85 (2013). Deidentifying data by way of only open sourcing spectra, electropherograms, and pattern comparisons may also be a means to open some areas of forensic research. The PCAST Report, *supra* note 2 at 103 also suggested using the identifying information of deceased individuals in foundational research to overcome some privacy issues.

²⁰⁹ See sources at *supra* note 42.

forensic science—has sought to improve confidence in its product, the answer has often been through open justice: opening courtrooms, permitting media scrutiny, and publishing decisions. It may be time that forensic science follows suit.

APPENDIX A: JOURNAL REVIEW METHODOLOGY

Table 1 contains our review of the publication guidelines of forensic science journals. Our full results can be found online.²¹⁰ We do not intend this to serve as an authoritative or exhaustive list of journals. Rather, we attempted to compile a snapshot of publication standards at journals that publish about forensic science (see the search methodology below). We did not include guild journals (eg the journal of the Association of Firearm and Toolmark Examiners), which may bias our review such that the state of forensic publishing appears more open.²¹¹ The results are current as of October 18, 2018.

Search Methodology

We began by including the journals listed under the subcategory ‘Medicine, Legal’ on the Web of Science.²¹² The Web of Science describes this subcategory as follows:

Medicine, Legal covers resources on all aspects of medical legal issues, including government regulations and policies, malpractice, toxicological and pharmacological regulations, clinical therapeutic patents and other critical legal issues at the interface of law, medicine, and healthcare. The category also covers resources dealing with the various branches of forensic science.

This subcategory included 16 journals: Forensic Science International-Genetics, Regulatory Toxicology and Pharmacology, International Journal of Legal Medicine, Forensic Science Medicine and Pathology, Forensic Science International, Science & Justice, Legal Medicine, Journal of Forensic Sciences, Journal of Forensic and Legal Medicine, Medical Law Review, Journal of Law Medicine & Ethics, Australian Journal of Forensic Sciences, American Journal of Forensic Medicine and Pathology, Rechtsmedizin, Medicine Science and the Law, and Romanian Journal of Legal Medicine.

We then supplemented the Web of Science results with our own internet searches, using combinations of terms including ‘forensic’, ‘science’, ‘journal’, ‘publication’, and ‘peer review(ed)’. This search returned 14 additional journals, for a total of 30. We found impact factors either found on the journal’s website or through the Web of Science’s listing (which purports to be accurate as to 2017).

Coding Methodology

We determined if the journal was a TOP signatory by consulting the Open Science Framework’s list of TOP signatories.²¹³

²¹⁰ Authors, *Openness of Forensic Journals Updated Nov 9 2018*, <https://osf.io/Spk7j/> (last accessed 2019). For what may be seen as a companion to this study, however one focused on forensic practice (versus research), see Cásarez & Thompson, *supra* note 2, at 1027-1030.

²¹¹ Ballantyne et al., *supra* note 177, at 70.

²¹² See Web of Science, <http://login.webofknowledge.com> (last accessed 2019).

²¹³ Center for Open Science, *supra* note 98.

Information about whether the journal is open access, its author publication fees, and whether the journal had implemented (formally or informally) TOP standards was determined through material on the journals' websites.

No journals had formally implemented TOP. Still, some included language encouraging some open practices that coincided with TOP standards on their websites. If such language was present, we coded them as 'Not Implemented: Encouraged' on the full spreadsheet and as 'TOP 0/Enc' on Table 1. If no such encouraging language was present, we coded them as 'Not Implemented: Says Nothing' on the full spreadsheet and as '0' on Table 1.

As to open access, language on journal websites varied considerably. We settled on the following coding scheme:

- 'Open access only': the journal only provides only an open access option.
- 'Open access by default': the default choice for authors is open access, but they can opt for subscription-based publishing.
- 'Choice of open or subscription': there is no default, the author chooses open or subscription based.
- 'Subscription by default': subscription is the default choice, but authors can opt for open access.
- 'No open access option': the journal does not provide any open access option.