Improving Expert Evidence: The Role of Open Science and Transparency

Jason M. Chin, Bethany Growns & David T. Mellor

BOTH SCIENCE AND EXPERT evidence law are undergoing significant changes. In this article, the authors compare these two movements—the open science movement and the evidence-based evidence movement. The open science movement encompasses the recent discovery of many irreproducible findings in science and the subsequent move towards more transparent methods. The evidence-based evidence movement is the discovery that many forms of expert evidence are unreliable, and that they have contributed to wrongful convictions. The authors identify similarities between these movements, which suggest how courts and legal actors may learn from the open science movement to produce more accurate results. Expert witnesses should comport themselves as rigorous open scientists to produce evidence that is more susceptible to evaluation. Parties should be subjected to more specific and rigorous disclosure requirements because research has shown that even leading scientists find it easy to discount and suppress findings that do not support their hypotheses. And trial judges, as gatekeepers, should not defer to the generally accepted practices that have proven insufficient in the mainstream sciences.

LE DROIT DE LA preuve se basant sur la science et le droit de la preuve se basant sur la preuve d'expert ou d'experte subissent tous les deux des changements importants. Dans cet article, les auteurs et auteures comparent ces deux mouvements – le mouvement de la science ouverte et le mouvement de la preuve basée sur les preuves. Le mouvement de la science ouverte englobe la découverte récente de nombreuses conclusions scientifiques non reproductibles et le changement éventuel vers des méthodes plus transparentes. Le mouvement de la preuve fondée sur les preuves repose sur la découverte du fait que de nombreuses sortes de preuves ne sont pas fiables et que celles-ci ont mené à des condamnations injustifiées. Les auteurs et auteures soulèvent les similitudes entre ces deux mouvements, qui suggèrent comment les tribunaux et les intervenants et intervenantes du système judiciaire peuvent s'inspirer du mouvement de la science ouverte afin de pouvoir arriver à des résultats plus précis. Les témoins experts et témoins expertes devront se comporter comme des scientifiques méticuleux et méticuleuses afin de fournir des preuves plus susceptibles d'être évaluées. Les deux partis devront être soumis à des exigences de divulgation plus rigoureuses et plus précises parce que les recherches démontrent que même les scientifiques les plus réputés et réputées trouvent qu'il est très facile d'écarter ou de cacher les résultats qui

ne corroborent pas leurs hypothèses. Et les juges et juges du procès, en tant que gardiens et gardiennes, ne devraient pas se tourner vers des pratiques couramment acceptées qui se sont révélées insuffisantes dans le cadre des sciences de tous les jours.

CONTENTS

Improving Expert Evidence: The Role of Open Science and Transparency

Jason M. Chin, Bethany Growns & David T. Mellor[*]

I. INTRODUCTION

The way in which science is conducted and communicated is fundamentally changing. Scientists and journals are increasingly adopting practices aimed at making science more transparent, reproducible, and democratic.[1] This article will demonstrate several parallels between this movement in science—the open science movement—and similar trends in expert evidence law. In particular, the genesis of many aspects of the open science movement was the realization that longstanding practices had failed, allowing spurious findings to reach general acceptance.[2] A similar pattern

* Jason Chin (JD, Toronto; PhD, UBC) is a Lecturer at Sydney Law School. Bethany Growns (PhD/M.Psyc, UNSW) is a Post-Doctoral Fellow at Arizona State University. David Mellor (PhD, Rutgers) is the Director of Policy Initiatives at the Center for Open Science in Virginia, USA. We are indebted to Shari Diamond, Gary Edmond, Jeremy Freese, Glenn Harrell, Sarah Lawsky, and John Monahan for their insightful comments on previous drafts. We also thank the participants of the Second Annual Junior Faculty Forum for Law and STEM (Northwestern Pritzker School of Law) for helpful feedback. Many of the ideas herein were also discussed at the annual meeting of the Evidence-Based Forensics Initiative. Sarah Hamid provided indefatigable editorial and research support.

1 See The National Academies of Sciences, Engineering, and Medicine: Committee on Toward an Open Science Enterprise Board on Research Data and Information Policy and Global Affairs, *Open Science by Design: Realizing A Vision for 21st Century Research* (Washington, DC: The National Academies Press, 2018) [NASEM, "Open Science Report"]; Marcus R Munafò et al, "A Manifesto for Reproducible Science" (2017) 1 Nature Human Behaviour 1.

2 See Leif D Nelson, Joseph Simmons & Uri Simonsohn, "Psychology's Renaissance" (2018) 69 Annual Rev Psychology 511 at 512 [Nelson, Simmons & Simonsohn, "Psychology's Renaissance"]. NASEM, "Open Science Report", *supra* note 1 at 31.

has been observed in several classic fields of expert evidence.[3] These parallels have significant consequences for law—a field where flaws in its truth-determining mechanisms have contributed to grave miscarriages of justice.[4] This article's central thesis is that open science-inspired reforms align with, and further, the ideals of expert evidence: these reforms help produce knowledge that is susceptible to critical evaluation.

The open science movement responded to the discovery of several results, many previously seen as robust and well-established, that could not be independently reproduced.[5] While these examples could at first be disregarded as outliers, the scientific community has come to acknowledge that they reflect an endemic problem. For instance, large-scale attempts to reproduce established social scientific findings have only succeeded about 40–60% of the time and have reported considerably weaker findings.[6] Reflecting these surprising (non)findings, a survey of 1,576 scientists in the journal, *Nature*, reported that 90% of those surveyed believed that science had a reproducibility problem.[7] Over 50% of researchers reported having failed to reproduce another's finding.[8]

These meta-scientific revelations have begun to inform and inspire reform. For simplicity, we use the term "open science movement" to refer

---

3 See Michael J Saks & David L Faigman, "Failed Forensics: How Forensic Science Lost Its Way and How It Might Yet Find It" (2008) 4 Annual Rev L & Soc Science 149; Lisa Dufraimont, "New Challenges for the Gatekeeper: The Evolving Law on Expert Evidence in Criminal Cases" (2012) 58:3&4 Crim LQ 531; Gary Edmond & Kent Roach, "A Contextual Approach to the Admissibility of the State's Forensic Science and Medical Evidence" (2011) 61:3 UTLJ 343; Alan D Gold, *Expert Evidence in Criminal Law: The Scientific Approach*, 2nd ed (Toronto: Irwin Law, 2009).

4 See Brandon L Garrett & Peter J Neufeld, "Invalid Forensic Science Testimony and Wrongful Convictions" (2009) 95:1 Va L Rev 1; Ontario Ministry of the Attorney General, *The Commission on Proceedings Involving Guy Paul Morin: Report*, vol 1 (Toronto: Queen's Printer for Ontario, 1998) (Hon Fred Kaufman) [Kaufman Report]; Ontario Ministry of the Attorney General, *Inquiry into Pediatric Forensic Pathology in Ontario: Report*, vols 1–4 (Toronto: Queen's Printer for Ontario, 2008) (Hon Stephen T Goudge) [Goudge Report].

5 NASEM, "Open Science Report", *supra* note 1 at 31.

6 See e.g. Open Science Collaboration, "Estimating the Reproducibility of Psychological Science" (2015) 349:6251 Science 943 [OSC]; Colin F Camerer et al, "Evaluating the Replicability of Social Science Experiments in *Nature* and *Science* Between 2010 and 2015" (2018) 2 Nature Human Behaviour 637; US, Andrew C Chang & Phillip Li, *Is Economics Research Replicable? Sixty Published Papers from Thirteen Journals Say 'Usually Not'* (Finance and Economics Discussion Series 2015-083) (Washington, DC: Board of Governors of the Federal Reserve System, 2015).

7 See Monya Baker, "1,500 Scientists Lift the Lid on Reproducibility" (2016) 533:7603 Nature 452 (52% classified it as a significant problem and 38% said it was a "slight crisis" at 452).

8 *Ibid* at 453.

to the totality of these developments. The movement, however, is broader than what we will focus on in this article, with an earlier (but ongoing) part of the campaign focused on access to scientific articles (*i.e.* removing paywalls to publicly-funded research). Rather, we will devote most of our analysis to transparency and openness as ways to improve the rigour of research methods by reducing undisclosed flexibility. This component of the movement has taken on various names in the literature, such as the "replicability"[9] crisis and the "credibility revolution."[10]

In parallel with the open science movement, the forensic sciences (*i.e.* science designed to answer legal questions) and scientific evidence in court more broadly have also been subjected to increased scrutiny.[11] Although open science and evidence law have almost never been explicitly linked,[12] many of the issues are remarkably similar. For instance, the challenges in science flow from cognitive biases that focus scientists on the data that confirm their hypotheses at the expense of those that do not (despite both sets of data having equal evidential value).[13] These are the very biases present in expert evidence that have been uncovered by academics,[14] and reports from peak bodies of scientists and jurists convened to address the failures of the criminal justice system.[15] Moreover, both

9 See Harold Pashler & Eric-Jan Wagenmakers, "Editors' Introduction to the Special Section on Replicability in Psychological Science: A Crisis of Confidence?" (2012) 7:6 Perspectives on Psychological Science 528.

10 See Simine Vazire, "Implications of the Credibility Revolution for Productivity, Creativity, and Progress" (2018) 13:4 Perspectives on Psychological Science 411 [Vazire, "Credibility Revolution"].

11 Saks & Faigman, *supra* note 3.

12 But see Jason M Chin, "Psychological Science's Replicability Crisis and What it Means for Science in the Courtroom" (2014) 20:3 Psychol Pub Pol'y & L 225 [Chin, "Replicability Crisis"]; Chuan-Peng Hu et al, "Open Science As a Better Gatekeeper for Science and Society: A Perspective From Neurolaw" (2018) 63:23 Science Bulletin 1529.

13 Munafò et al, *supra* note 1 at 1.

14 See D Michael Risinger et al, "The *Daubert*/*Kumho* Implications of Observer Effects in Forensic Science: Hidden Problems of Expectation and Suggestion" (2002) 90:1 Cal L Rev 1; Gary Edmond et al, "Contextual Bias and Cross-Contamination in the Forensic Sciences: The Corrosive Implications for Investigations, Plea Bargains, Trials and Appeals" (2015) 14:1 L, Probability & Risk 1 [Edmond et al, "Contextual Bias"]; Gary Edmond & Emma Cunliffe, "Cinderella Story? The Social Production of a Forensic 'Science'" (2017) 106:2 J Crim L & Criminology 219 at 229, 244; Itiel E Dror & David Charlton, "Why Experts Make Errors" (2006) 56:4 J Forensic Identification 600 at 610; Itiel E Dror, David Charlton & Ailsa E Péron, "Contextual Information Renders Experts Vulnerable to Making Erroneous Identifications" (2006) 156:1 Forensic Science Intl 74.

15 See e.g. US, President's Council of Advisors on Science and Technology, *Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods* (Washington,

movements are associated with a lack of transparency[16] and a preoccupation with eminence over methodology.[17]

It is meaningful that mainstream scientists are being accused of many of the same practices that have resulted in wrongful convictions in law. Specifically, it means that it is insufficient that expert witnesses be directed to behave like scientists. Instead, they should behave like open scientists. In other words, expert witnesses should be expected to behave as *scientists should be expected to behave*: candidly sharing the results of research, avoiding appeals to status, and skeptically scrutinizing their own work and that of others.[18] By embracing these norms (rather than simply mainstream ones), expert witnesses can provide evidence that is both more trustworthy and more susceptible to rational evaluation.[19]

Parts II and III will go on to describe the geneses of the open science movement and the evidence-based evidence movement, respectively. Part IV will analyze these movements, identifying six points of comparison. These similarities suggest mutually applicable reforms—a commitment to transparency and openness can improve the accuracy of both science and expert evidence. Part V then delves into legal reforms. Part VI concludes with two ancillary benefits of open expert evidence: improved trust and efficiency.

Before delving into the substance of our article, we should provide a brief caution. While we will suggest that insights from the open science movement have much to offer fact-finding in court, we note that science and law do not share all of their values. Importantly, law must balance other interests, like procedural fairness, adversarial imbalance, and

---

DC: Executive Office of the President, 2016) at 32 [PCAST Report]; National Research Council of the National Academies, *Strengthening Forensic Science in the United States: A Path Forward* (Washington, DC: National Academies Press, 2009) at 184–85 [NRC Report]. Goudge Report, *supra* note 4 at 43, 69, 79, 153–56, 374–77.

16 See BA Nosek et al, "Promoting an Open Research Culture" (2015) 348:6242 Science 1422 at 1422 [Nosek et al, "TOP"] (in science). See also Gary Edmond et al, "Model Forensic Science" (2016) 48:5 Australian J Forensic Sciences 496 at 497–99 [Edmond et al, "Model Forensic Science"] (in law).

17 See Robert K Merton, *The Sociology of Science: Theoretical and Empirical Investigations* (Chicago: University of Chicago Press, 1973) at xvii; Simine Vazire, "Our Obsession with Eminence Warps Research" (2017) 547:7661 Nature 7 (in science). See also *R v Béland*, [1987] 2 SCR 398 at para 78, 43 DLR (4th) 641 [*Béland*]; *R v Mohan*, [1994] 2 SCR 9 at para 23, 114 DLR (4th) 419 [*Mohan*]; *Nguyen v R*, [2017] NSWCCA 4 at para 28 (in law).

18 Merton, *supra* note 17 at 269.

19 See generally Gary Edmond, "Forensic Science Evidence and the Conditions for Rational (Jury) Evaluation" (2015) 39:1 Melbourne UL Rev 77; Alan W Mewett & Peter J Sankoff, *Witnesses* (Toronto: Carswell, 2018) at 16.3.

finality.[20] For instance, courts should be sensitive to the fact that a well-heeled corporate defendant facing a product liability claim would often be expected to have access to more sophisticated, case-relevant scientific evidence than the plaintiff. On the other hand, the criminally accused often face resource constraints, making it difficult to hire a rebuttal expert. As a result, we do not suggest holding all parties to the highest standard of open scientific evidence.

Rather, as explained in Parts V and VI, openness of the foundations and application of expert knowledge simply results in evidence that is more susceptible to rational evaluation. Indeed, as noted above, the open science movement is sometimes referred to as the "credibility revolution"[21] because the reforms transcend value-laden categorizations of science and non-science, applying across fields of knowledge generation. Similarly, in law, academics have criticized rules that require slippery taxonomies of expert evidence (*e.g.* those that would give different scrutiny to science as opposed to what a court might characterize as non-science, or put less weight on evidence simply because it was generated for the purposes of litigation).[22] Openness, as we will demonstrate, cuts across quantitative and qualitative disciplines. Put simply, and as the meta-scientific findings demonstrate below, openness makes the expert opinion's strengths and weaknesses more apparent and can therefore promote justice.

II. FROM CRISIS TO RENAISSANCE: A BRIEF HISTORY OF THE OPEN SCIENCE MOVEMENT

> If a team of research psychologists were to emerge today from a 7-year hibernation, they would not recognize their field. Authors voluntarily posting their data. Top journals routinely publishing replication attempts, both failures and successes. Hundreds of researchers preregistering their studies. Crowded methods symposia at many conferences. Enormous

20 See generally Gary Edmond & Mehera San Roque, "Just(,) Quick and Cheap? Contemporary Approaches to the Management of Expert Evidence" in Michael Legg, ed, *Resolving Civil Disputes* (LexisNexis, 2016).

21 Vazire, "Credibility Revolution", *supra* note 10 at 411.

22 See Jason M Chin, "*Abbey* Road: The (Ongoing) Journey to Reliable Expert Evidence" (2018) 96:3 Can Bar Rev 422 at 424, n 2 [Chin, "*Abbey* Road"]. See also Susan Haack, "What's Wrong with Litigation-Driven Science? An Essay in Legal Epistemology" (2008) 38 Seton Hall L Rev 1053 at 1077 ("The fact that research is litigation-driven in the stronger sense, I have argued, makes it likely to be biased. Biased research doesn't necessarily produce false results; nor does it necessarily produce false results more often than true.").

> increases in sample sizes. Some top journals requiring the full disclosure of measures, conditions, exclusions, and the rules for determining sample sizes. Several multilab replication efforts accepted for publication before any data were collected. Overall, an unprecedented focus on replicability. What on earth just happened?[23]

The open science movement was spurred by a surprising number of reports of published studies proving to be irreproducible (another thrust of open science is dedicated to making paywalled scientific journals available to the public, especially when the underlying research was publicly funded).[24] In other words, researchers attempted to recreate the findings of previous studies, but found inconsistent or considerably smaller effects. Such incidences were concentrated in pre-clinical and clinical medical research[25] and psychology.[26] Many fields, however, are struggling with the reproducibility of their findings.[27] These failures to reproduce results inspired large-scale systematic studies (mentioned in Part I), finding that studies published in eminent journals regularly proved irreproducible.[28]

These demonstrations of systemic problems within science raised difficult questions. Most fundamentally, why were these studies, which carried the *indicia* of good science (*e.g.* testing with a low reported error rate and publication in leading journals)[29] nevertheless false positive findings?

---

23 Nelson, Simmons & Simonsohn, "Psychology's Renaissance", *supra* note 2 at 512.

24 NASEM, "Open Science Report", *supra* note 1 at 31–37.

25 See e.g. Francis S Collins & Lawrence A Tabak, "NIH Plans to Enhance Reproducibility" (2014) 505:7485 Nature 612; Leonard P Freedman, Iain M Cockburn & Timothy S Simcoe, "The Economics of Reproducibility in Preclinical Research" (2015) 13:6 PLoS Biology 1; John PA Ioannidis, "Contradicted and Initially Stronger Effects in Highly Cited Clinical Research" (2005) 294:2 J American Medical Assoc 218; C Glenn Begley & Lee M Ellis, "Raise Standards for Preclinical Cancer Research" (2012) 483:7391 Nature 531.

26 Nelson, Simmons & Simonsohn, "Psychology's Renaissance", *supra* note 2 at 512–14.

27 See Denes Szucs & John PA Ioannidis, "Empirical Assessment of Published Effect Sizes and Power in the Recent Cognitive Neuroscience and Psychology Literature" (2017) 15:3 PLoS Biology 1 (in cognitive neuroscience). See also Chang & Li, *supra* note 6 (in economics); Daiping Wang et al, "Irreproducible Text-Book 'Knowledge': The Effects of Color Bands on Zebra Finch Fitness" (2018) 72:4 Evolution 961 (in evolutionary biology); John PA Ioannidis et al, "Replication Validity of Genetic Association Studies" (2001) 29:3 Nature Genetics 306 (in genetic association studies); Jeremy Freese & David Peterson, "Replication in Social Science" (2017) 43:1 Annual Rev Sociology 147 (sociologists are also currently wresting with reforms in their field).

28 See e.g. OSC, *supra* note 6; Camerer et al, *supra* note 6; Chang & Li, *supra* note 6.

29 See US, *Daubert v Merrell Dow Pharmaceuticals, Inc*, 509 US 579 (1993) at 593–595, 113 S Ct 2786 (1993) [*Daubert* 1993] (these are indeed three factors of scientific validity found in the US Supreme Court's foundational expert evidence decisions, as discussed in Part III).

Here, we note that many of these discoveries of false positives were not entirely surprising to some, in particular, sociologists of science, who long warned that science was more human and error prone than most realized.[30] Importantly, however, large-scale replication attempts (in the past, replications of previous work were rare) gave teeth to the sociological concerns and encouraged rigorous meta-scientific research into the precise weaknesses in science.[31]

One of the most widely-studied of those weaknesses is what has come to be known as "researcher degrees of freedom" or "questionable research practices" (QRPs).[32] These are *undisclosed* choices that researchers can use to increase their chances of finding a result that meets conventional levels of statistical certainty: "In the course of collecting and analyzing data, researchers have many decisions to make: Should more data be collected? Should some observations be excluded? Which conditions should be combined and which ones compared? Which control variables should be considered?"[33]

To understand how such tactics might give a reported study a misleading sheen, imagine if a basketball team could, instead of respecting a predetermined time limit, strategically decide to stop a match when they were winning—that would bias the game in their favour. Indeed, in one influential study, researchers found that use of just four[34] of these researcher degrees of freedom could inflate what would appear to be a 5% false positive rate to 60.7%.[35] Studies examining the damage QRPs can do were lent additional force by survey studies in psychology, ecology, and evolutionary biology, finding that researchers in those fields (responding anonymously)

30 See Simon A Cole & Alyse Bertenthal, "Science, Technology, Society, and Law" (2017) 13:1 Annual Rev L & Soc Science 351 at 353.

31 Matthew C Makel, Jonathan A Plucker & Boyd Hegarty, "Replications in Psychology Research: How Often Do They Really Occur?" (2012) 7:6 Perspectives on Psychological Science 537; Brian A Nosek & Timothy M Errington, "Making Sense of Replications" (2017) 6 eLife 1.

32 See Joseph P Simmons, Leif D Nelson, & Uri Simonsohn, "False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything As Significant" (2011) 22:11 Psychological Science 1359 at 1359, 1361 [Simmons, Nelson & Simonsohn, "False-Positive Psychology"].

33 *Ibid* at 1359.

34 *Ibid* at 1360 (the four researcher degrees of freedom included "flexibility in (a) choosing among dependent variables, (b) choosing sample size, (c) using covariates, and (d) reporting subsets of experimental conditions.").

35 *Ibid* at 1359, 1361.

used some questionable research practices at a high rate (50–60%, in some cases).[36]

Reflecting the centrality of openness to scientific progress, the National Academies of Sciences, Engineering, and Medicine (NASEM) issued a report in 2018 laying out a plan to improve science.[37] The report acknowledged that "research conducted openly and transparently leads to *better science*" and that scientific findings "are more likely to be credible—or found wanting—when they can be reviewed, critiqued, extended, and reproduced by others."[38] We will now briefly detail these reforms as they have been expressed in peer review and publication guidelines, scientific methodology, as well as in key enhancements to the infrastructure of science (*e.g.* a central website to store data and promote collaboration, as discussed below).

Reforming practices at academic journals is crucial because they are one of the primary venues by which science is vetted and transmitted. Indeed, the journal, *Nature*, expressly admitted its own role in the replicability crisis: "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 information for other researchers to assess results properly."[39] *Nature* went on to institute a host of improvements to its editorial policy, including enhanced reporting of methodology and abolishing space restrictions in those sections.[40] In medical research, these practices have been linked with modest improvements in reporting practices.[41]

More recently, a committee of researchers, journal editors, and funding organizations devised the Transparency and Openness Promotion (TOP)

36 See e.g. Leslie K John, George Loewenstein & Drazen Prelec, "Measuring the Prevalence of Questionable Research Practices with Incentives for Truth Telling" (2012) 23:5 Psychological Science 524; Hannah Fraser et al, "Questionable Research Practices in Ecology and Evolution" (2018) 13:7 PLoS ONE 1.

37 NASEM, "Open Science Report", *supra* note 1.

38 *Ibid* at 107 [emphasis in original].

39 See Editorial, "Announcement: Reducing Our Irreproducibility" (2013) 496:7446 Nature 398 at 398.

40 *Ibid*.

41 See Lucy Turner et al, "Does Use of the CONSORT Statement Impact the Completeness of Reporting of Randomised Controlled Trials Published in Medical Journals? A Cochrane Review" (2012) 1:60 Systematic Reviews 1; SeungHye Han et al, "A Checklist is Associated with Increased Quality of Reporting Preclinical Biomedical Research: A Systematic Review" (2017) 12:9 PLoS ONE.

Guidelines.[42] They include eight standards that can be implemented at three levels of rigour, from encouragement to requiring that the standard be met (with journal verification that it was met).[43] Several standards focus on transparency, in particular that of the author's data, analysis, materials, and research design. The standards also provide for replication (*i.e.* studies aimed at directly recreating previous studies) to address the barriers to funding and publishing important confirmatory research.

Further, the guidelines include a preregistration standard. Preregistration guards against QRPs by requiring that researchers record their procedures *before* collecting data.[44] As discussed below, the preregistration record is uneditable and often housed in an online database. Such precautions prevent researchers from strategically suppressing measures that did not support their hypothesis. This is not to say that deviations from the preregistered plan are not allowed and that the data collected from studies with deviations are worthless. Rather, it simply ensures that the peer reviewer and the consumer of the science know that the plan changed, so that such deviations can be evaluated.

As of February 2009, nearly 5,000 journals have endorsed the TOP Guidelines and committed to review their own standards within a year.[45] Many have implemented the guidelines at varying levels.[46] For instance, the journal, *Science*, recently revised its editorial policy to require that authors make their data available, subject to "truly exceptional circumstances."[47]

The changes in journals are scaffolded by new tools available to researchers.[48] Notably, the Open Science Framework (OSF) is a free web

---

42 Nosek et al, "TOP", *supra* note 16.

43 *Ibid* at 1424.

44 Nelson, Simmons & Simonsohn, "Psychology's Renaissance", *supra* note 2 at 519–20. See also David Mellor, Simine Vazire & D Stephen Lindsay, "Transparent Science: A More Credible, Reproducible, and Publishable Way to Do Science" in Robert J Sternberg, ed, *Guide to Publishing in Psychology Journals*, 2nd ed (Cambridge: Cambridge University Press, 2019); Brian A Nosek et al, "The Preregistration Revolution" (2018) 115:11 Proceedings National Academy Sciences 2600 [Nosek et al, "Preregistration"].

45 See Center for Open Science, "Current Signatories," online: *Center for Open Science* <cos.io/our-services/top-guidelines/> (the list of journals and publishers on file with the Center for Open Science). See also Center for Open Science, "Transparency and Openness Promotion (TOP) Guidelines", online: *Open Science Framework* <osf.io/pvf56/> (listing the obligations involved with endorsing the guidelines).

46 See Center for Open Science, "Implementing TOP," online: *Center for Open Science* <cos.io/our-services/top-guidelines/>.

47 See Science, "*Science* Journals: Editorial Policies," online: *Science* <www.sciencemag.org/authors/science-editorial-policies>.

48 See generally Munafò et al, *supra* note 1 at 2–7 (for a review).

platform for open science.[49] It provides support and infrastructure for users at all stages of the research process. For instance, it allows researchers to conform with TOP Guidelines by preregistering their study or by sharing data, code, and other digital materials. Such sharing is important because data analysis and the computation underlying research is increasingly complex and central to the scientific process. As a result, sharing can produce efficiencies and the chance that researchers will catch each other's mistakes.[50]

III. THE EVIDENCE-BASED EVIDENCE MOVEMENT

As with the open science movement, the evidence-based evidence movement has been informed by past mistakes.[51] Most notably, legal scholars have documented the numerous miscarriages of justice attributable to expert witnesses giving invalid scientific evidence.[52] For instance, one US study found invalid forensic science in 63% of cases in which forensic scientific testimony was tendered.[53] These revelations lent credence to long-standing worries that the practice of forensic science was subject to a host of errors and uncertainties. Throughout this period, the rules regulating the admission of scientific evidence were becoming formally more demanding (spurred by the admission of junk science in US civil litigation).[54] In this part, we will briefly review these three ingredients: wrongful convictions, academic research focused on the forensic sciences, and the legal regulation of scientific evidence.

---

49 See Center for Open Science, "Open Science Framework," online: *Open Science Framework* <osf.io>; Munafò et al, *supra* note 1 at 4. See also "As Predicted," online: *As Predicted* <aspredicted.org> (another example of a free web platform for preregistration and open science).

50 See generally Victoria Stodden, "Trust Your Science? Open Your Data and Code", *Amstat News* 409 (July 2011) 21, online (pdf): <magazine.amstat.org/wp-content/uploads/2011an/July2011.pdf>.

51 See David Paciocco, "Taking A 'Goudge' Out of Bluster and Blarney: An 'Evidence-Based Approach' to Expert Testimony" (2009) 13:2 Can Crim L Rev 135 (by using the term "evidence-based evidence movement," we are attempting to describe a general trend across several, sometimes related, areas of legal scholarship) [Paciocco, "Expert Testimony"]; D Michael Risinger, "Navigating Expert Reliability: Are Criminal Standards of Certainty Being Left on the Dock?" (2000) 64:1 Alb L Rev 99.

52 Garrett & Neufeld, *supra* note 4; Kaufman Report, *supra* note 4; Goudge Report, *supra* note 4.

53 Garrett & Neufeld, *supra* note 4 at 14, 29.

54 *Daubert* 1993, *supra* note 29. See also *General Electric Co v Joiner*, 522 US 136 (1997), 118 S Ct 512 (1997); *Kumho Tire Co, Ltd v Carmichael*, 526 US 137 (1999), 119 S Ct 1167 (1999) [*Kumho*].

As with the US, several Canadian wrongful convictions have been caused—at least in part—by invalid or misleading forensic science. For instance, Guy Paul Morin's wrongful conviction was, in part, based on scientifically invalid hair, fibre, and blood testing.[55] About a decade later, Justice Goudge's pivotal *Inquiry into Pediatric Forensic Pathology in Ontario* (the "Goudge Report") found many failings in the work of pediatric forensic pathologist Charles Smith. Smith's invalid testimony contributed to 14 wrongful convictions.[56] In the child protection arena, many children were taken from their parents—the "capital punishment"[57] of child welfare law—based on invalid hair tests from the "Motherisk" program that purported to detect the use of drugs or alcohol in parents. These failings were detailed in the recent reports of Justice Lang and Justice Beaman.[58]

Wrongful convictions enlivened an existing scholarship evaluating the validity and reliability of the forensic sciences, and their suitability as inculpatory evidence.[59] Much of this research had warned that many forensic practices (*e.g.* fingerprint examiners comparing a found print to a suspect's print) had never actually been tested and likely were more error-prone and subjective than practitioners were acknowledging in court. Moreover, even if these scientific foundations were established, forensic practitioners have long resisted practices designed to resist bias, such as blinding themselves (*i.e.* keeping themselves unaware) to the identity of the accused and other visceral case details irrelevant to their task.[60] These warnings eventually reached critical mass in the form of a review by the National Research Council of the National Academies of Science (the "NRC Report").[61] The NRC Report was quite frank in its criticism of the forensic sciences and the failure of courts to regulate such evidence: "in a number of forensic

55 Kaufman Report, *supra* note 4 at 12, 327.

56 Goudge Report, *supra* note 4. See also Emma Cunliffe & Gary Edmond, "What Have We Learned? Lessons from Wrongful Convictions in Canada" in Benjamin L Berger, Emma Cunliffe & James Stribopoulos, eds, *To Ensure that Justice Is Done: Essays in Memory of Marc Rosenberg* (Toronto: Thomson Reuters, 2017) 129 at 133.

57 See Ontario, *Harmful Impacts: The Reliance on Hair Testing in Child Protection: Report of the Motherisk Commission* (Toronto: Ministry of the Attorney General, 2018) (Hon Judith C Beaman) xxii–xxiii, quoting Justice Susan E Lang [Beaman Report].

58 Beaman Report, *ibid*. See also Ontario, Independent Reviewer, *Report of the Motherisk Hair Analysis Independent Review* (Toronto: Ministry of the Attorney General, 2015) (Hon Susan E Lang) [Lang Report].

59 See e.g. Jennifer L Mnookin, "The Courts, the NAS, and the Future of Forensic Science" (2010) 75:4 Brook L Rev 1209 at 1228–34 [Mnookin, "The Courts"].

60 Saks & Faigman, *supra* note 3 at 158. See also Edmond et al, "Contextual Bias", *supra* note 14.

61 NRC Report, *supra* note 15.

science disciplines, forensic science professionals have yet to establish either the validity of their approach or the accuracy of their conclusions, and the courts have been utterly ineffective in addressing this problem."[62] The NRC Report was followed by the 2016 (US) President's Council of Advisors on Science and Technology (PCAST) Report on forensic science, which found that little had changed since the earlier report.[63]

The third part of the evidence-based evidence story relates to the latter part of the NRC Report's conclusion—the (in)effectiveness of courts in regulating scientific evidence. During the early 1990s, courts in both the US and Canada shifted from a deferential approach to expert evidence to an approach that is—at least formally—more hands-on.[64] In the US, the foundational case is *Daubert*.[65] In *Daubert*, the Supreme Court overruled the previous doctrine that allowed expert evidence when it was generally accepted by the scientific community from which it came (this previous rule was established in *Frye v United States*).[66]

The Court in *Daubert* held that trial judges must not defer to general acceptance as they had in *Frye*, but must instead evaluate the science themselves to determine if it is sufficiently reliable to admit into court.[67] To assist trial judges in exercising their newly enhanced gatekeeping responsibility, the Supreme Court provided four factors: (1) whether and how the opinion had been tested; (2) the peer review and publication status of the opinion; (3) the error rate associated with the opinion; and (4) its acceptance within the relevant field of knowledge.[68]

Canadian courts responded to the proliferation of expert evidence in a similar way and along roughly the same timeline. In 1993, the Supreme Court of Canada issued its decision in *R v Mohan*.[69] *Mohan*, which still stands as Canada's leading expert evidence decision, clarified and strengthened the rules for admitting expert opinion. Justice Sopinka, writing for the court, held that expert evidence is only admissible if it meets four criteria: (1) relevance; (2) necessity in assisting the trier of fact; (3) no other exclusionary rule applies; and (4) the tendered expert is properly qualified.[70]

62 *Ibid* at 53.

63 PCAST Report, *supra* note 15 at 122.

64 *Mohan*, *supra* note 17.

65 *Daubert* 1993, *supra* note 29.

66 293 F 1013 (1923), 34 ALR 145.

67 *Daubert* 1993, *supra* note 29 at 580, 589.

68 *Ibid* at 592–94.

69 *Mohan*, *supra* note 17. See also Sidney N Lederman, Alan W Bryant & Michelle K Fuerst, *The Law of Evidence in Canada*, 4th ed (Markham: LexisNexis, 2014) at 783–90.

70 *Mohan*, *supra* note 17 at 20.

Relevance, under *Mohan*, includes both logical relevance and a balancing of the benefits and costs of admitting the evidence (sometimes referred to as legal relevance).[71] These factors include the reliability of the evidence and the ability of the jury to rationally evaluate the basis of the opinion.[72] Furthermore, Justice Sopinka remarked that "expert evidence which advances a novel scientific theory or technique" should receive "special scrutiny," including meeting a threshold level of reliability and being essential to the trial.[73] Post-*Mohan* decisions elaborated on how reliability should be assessed,[74] with the Supreme Court expressly applying the *Daubert* factors in *R v J-LJ*.[75]

Several post-*R v J-LJ* developments should be noted. In 2015, the Supreme Court in *White Burgess* accepted and refined an appellate court's reformulation of the *Mohan* framework into a two-step test.[76] At the first step, the evidence's proponent must establish logical relevance, plus the three other threshold conditions from *Mohan*.[77] Also at the first stage, if the evidence is "based on novel or contested science or science used for a novel purpose,"[78] then it must be scientifically valid and reliable pursuant to *Daubert*.[79] At the second stage, the trial judge must consider legal relevance, a calculus that, as noted above, includes the reliability of the evidence.[80] *White Burgess* is also well known for establishing that an expert's lack of independence and impartiality may be cause to exclude that evidence (rather than simply being a matter of weight).[81] Recent high-profile decisions have relied on these new rules to exclude expert evidence.[82] But, while this new interest in excluding expert witnesses for partiality is

71 See *Mohan*, *supra* note 17. See also *White Burgess Langille Inman v Abbott and Haliburton Co*, 2015 SCC 23 at paras 23–24 [*White Burgess*] (for a summary of the *Mohan* analysis).

72 *Mohan*, *supra* note 17 at 21–22; *White Burgess*, *supra* note 71 at para 24.

73 *Mohan*, *supra* note 17 at 25. See also *R v Dimitrov* (2003), 68 OR (3d) 641 at para 37, 181 CCC (3d) 554.

74 Mewett & Sankoff, *supra* note 19 at 16.3(c)(i).

75 2000 SCC 51 at paras 33–35, [*JLJ*].

76 *White Burgess*, *supra* note 71 at paras 23–24.

77 *Ibid* at para 23 (the other threshold conditions are necessity, absence of an exclusionary rule, and a properly qualified expert).

78 *Ibid*.

79 *Ibid*.

80 *Ibid* at para 24.

81 *Ibid* at paras 46–54. See also Mewett & Sankoff, *supra* note 19 at 16.8(c) (for a lucid review and discussion of the pre-*White Burgess* law).

82 See e.g. *Bruff-Murphy v Gunawardena*, 2017 ONCA 502 at paras 42–70; *R v McManus*, 2017 ONCA 188 at paras 63–75; *JP v British Columbia (Children and Family Development)*, 2017 BCCA 308 at paras 221–39. See generally Emma Cunliffe, "A New Canadian Paradigm?

promising, other courts have entertained creative ways to circumvent the expert evidence rules. For instance, science is regularly characterized as "specialized knowledge" and thus, granted deference.[83]

In reviewing the Goudge Report, (now Justice) David Paciocco suggested it represented a paradigm shift. Judges should no longer allow experts to simply say "trust me" because that previously gave way to the many miscarriages of justice documented in the Report.[84] Instead, judges should insist that the expert "show me."[85] As we will detail in what follows, the "show me" approach is also the foundational prescription of open science reformists. Unfortunately, openness is still not regularly provided by experts, nor demanded by trial judges.

IV. SIX PARALLEL CHALLENGES

There are six commonalities shared by the open science movement and the evidence-based evidence movement: (a) flawed incentives, (b) excessive flexibility, (c) motivated reasoning, (d) immodesty, (e) appeals to eminence, and (f) some fraudulent practice. There are at least two reasons this comparison matters. First, it suggests shared remedies. The methods and principles that open science reformists have advocated for may inform how expert evidence ought to be produced and presented to courts. Second, these commonalities reinforce the notion that courts simply cannot defer to mainstream scientific norms and practices that have often proved inadequate. Rather, they should be aware of the reforms going on within science and, when appropriate, should hold experts to these new standards.

A. A Flawed Incentive System

> Publishing…lies at the very heart of modern academic science—at levels ranging from the epistemic certification of scientific thought to the more personal labyrinths of job security, quality of life, and self-esteem.[86]

Judicial Gatekeeping and the Reliability of Expert Evidence" in Paul Roberts & Michael Stockdale, eds, *Forensic Science Evidence and Expert Witness Testimony: Reliability Through Reform?* (Cheltenham, UK: Edward Elgar, 2018).

83 See Chin, "*Abbey* Road", *supra* note 22 at 424.

84 Paciocco, "Expert Testimony", *supra* note 51 at 146, 155–56.

85 *Ibid* at 156.

86 Michael J Mahoney, "Open Exchange and Epistemic Progress" (1985) 40:1 American Psychologist 29 at 30.

> Their livelihoods; their everyday relations with colleagues, superiors, and sometimes police; their personal identities—all are tied up in thinking and behaving as their group expects. And what is expected of them is to help complete the case against people suspected of having committed crimes.[87]

These two quotes, the first about academic scientists and the second about expert witnesses, are eerily similar. This is because both movements find their source in a misaligned incentive system.

In the open science movement, the culpable incentive is publication. Publication is the currency of the academic sciences, influencing promotion, grant success, and personal well-being.[88] Unfortunately, a finding's publishability is often a poor gauge of its truth value.[89] Journals value novelty over rigour.[90] The academic sciences have also not historically published replication research (*i.e.* findings confirming or disconfirming previous studies).[91] Researchers seeking publication are aware of these rules and play the game. They conduct research with small sample sizes, do not attempt (or report) replications, and employ QRPs.[92] Results that did not uncover an interesting finding (*e.g.* the result was not statistically significant) sit in file drawers.[93] As a result, the strength of published findings is likely overstated.

Flawed incentives also contribute to the problems identified in the evidence-based evidence movement—they flow from an adversarial culture that incentivizes favourable legal results over truth.[94] For example, the prosecutors who decide how to deploy forensic scientific expertise have

87 Michael J Saks & Barbara A Spellman, *The Psychological Foundations of Evidence Law* (New York: New York University Press, 2016) at 209.

88 Mahoney, *supra* note 86.

89 See Brian A Nosek, Jeffrey R Spies & Matt Motyl, "Scientific Utopia: II. Restructuring Incentives and Practices to Promote Truth Over Publishability" (2012) 7:6 Perspectives on Psychological Science 615 [Nosek, Spies & Motyl, "Utopia"].

90 See Roger Giner-Sorolla, "Science or Art? How Aesthetic Standards Grease the Way Through the Publication Bottleneck But Undermine Science" (2012) 7:6 Perspectives on Psychological Science 562.

91 Makel, Plucker & Hegarty, *supra* note 31.

92 See Marjan Bakker, Annette van Dijk & Jelte M Wicherts, "The Rules of the Game Called Psychological Science" (2012) 7:6 Perspectives on Psychological Science 543.

93 See Kay Dickersin, "The Existence of Publication Bias and Risk Factors for Its Occurrence" (1990) 263:10 J American Medical Assoc 1385 at 1386.

94 Saks & Spellman, *supra* note 87 at 208–09. See also David M Paciocco, "Unplugging Jukebox Testimony in an Adversarial System: Strategies for Changing the Tune on Partial Experts" (2009) 34:2 Queen's LJ 565 at 573–74 [Paciocco, "Jukebox"] (David Paciocco would describe the misaligned incentive problem as an actual or perceived lack of independence between the expert and trial process).

drawn criticism for seeking convictions instead of serving the administration of justice.[95] Such behaviour has sometimes taken the form of failing to disclose exculpatory evidence, as in the miscarriages of justice in the convictions of Roméo Phillion[96] and Donald Marshall Jr.[97]

Expert witnesses are not immune from biases flowing from the adversarial system.[98] Courts have frequently worried that expert testimony may be tinctured by the employment relationship with the party tendering them.[99] This "association bias" is heightened by "selection bias," the fact that experts are often chosen because they have a view favourable to that of the proffering party (David Paciocco refers to these two biases together as "adversarial bias").[100] The NRC Report, for instance, chronicled "significant concerns" with the independence of forensic scientists because they are often employed by the police.[101]

B. Excessive Flexibility

Misaligned incentives become a problem when there is flexibility to act on them. Open science researchers have documented "undisclosed flexibility" in the research process (see Part II) that gave way to the QRPs that many scientists used to artificially inflate the publishability of their findings.[102]

The same flexibility has contributed to issues with expert evidence. For example, the NRC Report found great flexibility in how forensic examiners report their findings.[103] It also found a general failure to establish

95 See e.g. Adam Benforado, *Unfair*: *The New Science of Criminal Injustice* (New York: Crown, 2015) at 27–29.

96 See *R v Phillion*, 2009 ONCA 202 [*Phillion*]. See also Innocence Canada, "Exonerations" (2019), online: *Innocence Canada* <www.aidwyc.org/cases/historical/romeo-phillion/>.

97 See Nova Scotia, *Royal Commission on the Donald Marshall, Jr, Prosecution: Digest of Findings and Recommendations 1989* (Halifax: McCurdy's Printing and Typesetting, 1989) [Donald Marshall Inquiry]. See also *R v Taillefer*, 2003 SCC 70 [*Taillefer*]; Garrett & Neufeld, *supra* note 4 (Garrett and Neufeld, in their empirical study of wrongful convictions in the US, found many cases in which the prosecution failed to disclose exculpatory forensics).

98 See Daniel C Murrie et al, "Are Forensic Experts Biased by the Side That Retained Them?" (2013) 24:10 Psychological Science 1889.

99 *White Burgess*, *supra* note 71 at para 11 citing *Abinger v Ashton* (1873), 17 LR Eq 358 at 374.

100 Paciocco, "Jukebox", *supra* note 94 at 575–81.

101 NRC Report, *supra* note 15 at 183–84. See also Garrett & Neufeld, *supra* note 4 at 13 (in the US, Garrett and Neufeld performed an empirical examination of wrongful convictions that included invalid forensic scientific evidence—the majority of experts in that study were employed by police crime labs).

102 Simmons, Nelson & Simonsohn, "False-Positive Psychology", *supra* note 32 at 1359–63.

103 See e.g. PCAST Report, *supra* note 15 at 83; NRC Report, *supra* note 15 at 185–186.

or commit to guidelines requiring that examiners be blind to potentially biasing information, such as the nature of the crime and the identity of the suspect.[104] In Canada, Justice Goudge singled out excessive flexibility as a contributor to the miscarriages of justice he reviewed: "Our systemic review of autopsy practices in Dr. Smith's years revealed the absence of any articulated principles…on which a set of best practices could be built."[105] Even validation studies (that scaffold the work of examiners) in forensic science have drawn criticism along similar lines to that in mainstream science. For instance, the PCAST Report noted that unclear exclusion rules in such studies (*e.g.* should an examiner be excluded for unusually low performance or could that be chalked up to a clerical error?) could bias their results.[106]

C. Motivated Reasoning

> [A] major challenge for scientists is to be open to new and important insights while simultaneously avoiding being misled by our tendency to see structure in randomness. The combination of apophenia (the tendency to see patterns in random data), confirmation bias (the tendency to focus on evidence that is in line with our expectations or favoured explanation), and hindsight bias (the tendency to see an event as having been predictable only after it has occurred) can easily lead us to false conclusions.[107]

Flawed incentives and excessive flexibility provide fertile ground for "motivated reasoning."[108] This term reflects the fact that our thought processes are not always (or often) rational, calculated, and transparent.[109] Rather, cognition is inherently and implicitly motivated; our conclusions are influenced by processes like contextual bias (*i.e.* the effect of cues in the environment on our reasoning)[110] and confirmation bias (*i.e.* the effect of pre-existing conclusions on our reasoning).[111] As these processes

104 PCAST Report, *supra* note 15 at 33.
105 Goudge Report, *supra* note 4 at 44.
106 PCAST Report, *supra* note 15 at 95.
107 Munafò et al, *supra* note 1 at 1.
108 See Ziva Kunda, "The Case for Motivated Reasoning" (1990) 108:3 Psychological Bulletin 480.
109 *Ibid*.
110 Edmond et al, "Contextual Bias", *supra* note 14. See also PCAST Report, *supra* note 15 at 31–32; Risinger et al, *supra* note 14 (Michael Risinger refers to these as "observer effects").
111 Goudge Report, *supra* note 4 at 424–25. See also Raymond S Nickerson, "Confirmation Bias: A Ubiquitous Phenomenon in Many Guises" (1998) 2:2 Rev General Psychology 175.

occur unconsciously, their operation is unimpeachable through cross and self-examination.[112]

As reflected in the quote that began this subsection, scientists long believed that the strictures of science protected them against these biasing forces.[113] They were deceiving themselves. The incentive to publish is strong and most editors did not always hold authors to very rigorous and transparent standards.[114] This gratuitous flexibility in the scientific method contributed to an often-irreproducible body of research.[115]

In law, motivated reasoning is widely considered to have contributed to numerous wrongful convictions and accusations.[116] Subsequent research systematically exposing forensic professionals to biasing information finds that their decisions are indeed influenced by these often-irrelevant details.[117] In response, leading bodies like PCAST and the NRC have encouraged forensic scientists to adopt blinding procedures.[118]

D. Epistemological Immodesty

The ubiquity of motivated reasoning should foster what some researchers have termed epistemological "humility"[119] and "modesty."[120] In other words, scientists and expert witnesses should be careful to not overstate their conclusions and to couch them in the appropriate levels of uncertainty. Unfortunately, both scientists and expert witnesses have not always upheld this ideal.

Within mainstream sciences, researchers long employed QRPs but refused to acknowledge the uncertainty they produced.[121] Somewhat ironically, this was even found in the work of a psychologist who won a Nobel

112 See Kathleen A Kennedy & Emily Pronin, "Bias Perception and the Spiral of Conflict" in Jon Hanson, ed, *Ideology, Psychology, and Law* (New York: Oxford University Press, 2012) at 411–414.

113 Munafò et al, *supra* note 1 at 1–2.

114 Giner-Sorolla, *supra* note 90.

115 Simmons, Nelson & Simonsohn, "False-Positive Psychology", *supra* note 32.

116 See e.g. PCAST Report, *supra* note 15 at 28; Goudge Report, *supra* note 4 at 153–56.

117 Dror & Charlton, *supra* note 14; PCAST Report, *supra* note 15 at 31.

118 PCAST Report, *supra* note 15 at 96; NRC Report, *supra* note 15 at 188.

119 NRC Report, *supra* note 15 at 106, 142. See also Jennifer L Mnookin, "The Validity of Latent Fingerprint Identification: Confessions of a Fingerprinting Moderate" (2008) 7:2 L, Probability & Risk 127 at 139.

120 Edmond et al, "Model Forensic Science", *supra* note 16 at 497–99.

121 Nelson, Simmons & Simonsohn, "Psychology's Renaissance", *supra* note 2 at 518.

Prize for studying motivated reasoning. Another scientist re-analyzed[122] the studies in a chapter of Daniel Kahneman's book, *Thinking, Fast and Slow*.[123] He found that the reproducibility of many of the studies Kahneman relied on was limited because they used small sample sizes. Acknowledging his previous immodesty, Kahneman replied: "I placed too much faith in underpowered studies...there is a special irony in my mistake because the first paper that Amos Tversky and I published was about the belief in the 'law of small numbers,' which allows researchers to trust the results of underpowered studies with unreasonably small samples."[124] Kahneman's self-scrutiny is laudable, but has typically been uncommon.

Similarly, in expert evidence, academics have frequently accused expert witnesses of immodesty. For instance, forensic bitemark[125] and fingerprint analysts[126] regularly stated—*in court*—that they could match a found pattern to the accused to the exclusion of all the world. There is no scientific or even logical basis for these claims.[127] These are unsupportable and decidedly immodest claims. Justice Goudge recognized similar brazenness in Charles Smith's testimony. Smith testified in a "dogmatic"[128] manner and regularly strayed outside the bounds of his expertise.[129]

E. A Preoccupation with Eminence

> The drive for eminence is inherently at odds with scientific values, and insufficient attention to this problem is partly responsible for the recent crisis of confidence in psychology and other sciences. Humans will always

122 See Ulrich Schimmack, Moritz Heene & Kamini Kesavan, "Reconstruction of a Train Wreck: How Priming Research Went Off the Rails" (2 February 2017), online (blog): *Replicability Index* <replicationindex.wordpress.com/2017/02/02/reconstruction-of-a-train-wreck-how-priming-research-went-of-the-rails/>.

123 See Daniel Kahneman, *Thinking, Fast and Slow* (Canada: Doubleday Canada, 2011).

124 See Alison McCook, "'I Placed Too Much Faith in Underpowered Studies:' Nobel Prize Winner Admits Mistakes" (20 February 2017), online (blog): *Retraction Watch* <retractionwatch.com/2017/02/20/placed-much-faith-underpowered-studies-nobel-prize-winner-admits-mistakes>.

125 See Michael J Saks et al, "Forensic Bitemark Identification: Weak Foundations, Exaggerated Claims" (2016) 3:3 J L & Biosciences 538 at 559–60.

126 Mnookin, "The Courts", *supra* note 59 at 1225–27.

127 *Ibid* (as Jennifer Mnookin noted, this also implies both that the source pattern is unique and that the analysts never make errors).

128 Goudge Report, *supra* note 4 at 16.

129 *Ibid* at 14.

> care about eminence. Scientific institutions and gatekeepers should be a bulwark against the corrupting influence of the drive for eminence.[130]

In the above quote, Simine Vazire implicates eminence in irreproducibility. This is because eminence bears a loose relationship with the actual truth value of scientific findings. But it is human nature to overweigh prestige.[131] Scientific safeguards have not always effectively controlled for eminence. For example, a 2017 study of research in computer science found that whether or not peer review was blind affected publication decisions (*i.e.* reviewers base their decisions, in part, on the work's author).[132]

Expert evidence shares this struggle with eminence. Before experts begin to provide their testimony to the factfinder, they are often led, by the tendering lawyer, through a lengthy review of their CV: impressive-sounding academic credentials, publications, journal editorships, and so on. This occurs unchallenged despite years of judicial admonishments against the danger that a lay jury will be unduly swayed by experts with impressive credentials.[133] Ironically, the factfinder is often not provided the most relevant *indicia* of "eminence"—the results of proficiency tests (*i.e.* is the expert actually good at the task at hand?).[134]

F. Intentional and Negligent Misbehaviour

To this point, we have primarily described the phenomenon of experts and scientists operating with too much flexibility and succumbing to motivated reasoning. But these are not the only types of errors out there. Rather, some are intentional.[135] Transparency and openness reforms should be designed to identify and deter not just exploitation of researcher degrees of freedom, but intentional acts as well.

---

130 See Simine Vazire, "Against Eminence" (11 April 2017) [unpublished, archived at PsyArXiv Preprints], online: <psyarxiv.com/djbcw>.

131 See Jon Hanson & David Yosifon, "The Situational Character: A Critical Realist Perspective on the Human Animal" (2004) 93:1 Geo L J 1 at 6–13.

132 See Andrew Tomkins, Min Zhang & William D Heavlin, "Reviewer Bias in Single- Versus Double-Blind Peer Review" (2017) 114:48 Proceedings National Academy Sciences 12708.

133 *Mohan*, *supra* note 17 at 24.

134 See Brandon L Garrett & Gregory Mitchell, "The Proficiency of Experts" (2018) 166:4 U Pa L Rev 901. PCAST Report, *supra* note 15 at 57–59; Mnookin, "The Courts", *supra* note 59 at 1224–25, 1235–36, 1268–75.

135 Saks & Faigman, *supra* note 3 at 159.

Estimating the prevalence of fraudulent research practices is difficult, but many researchers have attempted it.[136] Daniele Fanelli performed a meta-analysis (*i.e.* empirical review) of several such studies and found that two percent of scientists admitted to having "fabricated, falsified, or modified" their data.[137] We are unaware of similar surveys of expert witnesses and it is precarious to infer intentionality simply because an expert gave scientifically invalid testimony.[138] Still, cases of intentional falsification have occurred.[139]

V. OPEN SCIENCE LESSONS FOR LEGAL ACTORS

The primary takeaway from the above comparison is that, in the case of both mainstream science and expert evidence, individuals—influenced by misaligned incentives—took advantage of unreported flexibilities in their disciplines. Existing safeguards in both fields were poorly equipped to prevent misleading and invalid results from reaching orthodoxy. This similarity suggests that the problems found in expert evidence cannot simply be remedied by adherence to mainstream scientific methods and norms. Rather, legal reform should be attuned to the main response of open science reformists: transparency.[140] In science, new initiatives focus on transparently reporting flexibilities in the research process, such that peer-reviewers and the public can rationally evaluate the science. As it happens, rational evaluation of evidence is also a foundational principle of expert evidence law.[141] We will now suggest how open science may be embraced by legal actors: the experts, the advocates, and the gatekeepers.

136 See Daniele Fanelli, "How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data" (2009) 4:5 PLoS ONE 1.

137 *Ibid* at 8.

138 Garrett & Neufeld, *supra* note 4 at 76 ("Even with the benefit of bench notes or laboratory reports, one may not be able to ascertain whether experts falsified or concealed test results.").

139 Paciocco, "Jukebox", *supra* note 94 at 582–84.

140 See Edward K Cheng & G Alexander Nunn, "Beyond the Witness: Bringing A Process Perspective to Modern Evidence Law" 97:6 Tex L Rev [forthcoming in 2019] (this view also coincides with the focus on transparency of scientific processes proposed by Cheng & Nunn).

141 *Mohan*, *supra* note 17 at 24.

A. Adopting Open Scientific Practices: The Role of Expert Witnesses

> Where courts do not regulate the content of expert testimony, and defendants typically do not have experts with which to effectively counter State-proffered forensic testimony in criminal trials, the scientific standards within the forensic sciences are the most important source for regulating the content of forensic science testimony.[142]

The experts themselves are best placed to ensure their evidence is susceptible to rational evaluation. Our advice to them is straightforward: adopt open scientific practices (*e.g.* open data, preregistration, thorough reporting of methodology, and disclosure of flexibility in analysis). To make these recommendations more concrete, we will discuss them in the context of two recent controversial criminal law proceedings: one featuring expert social scientific evidence (*R v Abbey*), and one featuring expert forensic evidence (*R v Bornyk*).

In *Bornyk*, the forensic examiner's opinion exemplified many of the common criticisms of forensic science: he overclaimed and did not provide important knowledge to the court about the controversies in his field.[143] But, as we will see, the academic sociologist in *Abbey*, who followed mainstream norms for the most part, did not do much better. His methodology was laden with QRPs that were only revealed a decade after he first gave testimony. In both cases, open science may have assisted in more efficiently bringing these limitations to light.

In *Abbey*, the Crown proffered an expert sociologist named Mark Totten. He gave evidence indicating that the accused's teardrop tattoo meant he had killed a rival gang member.[144] This evidence was central to the Crown's theory that the accused shot the deceased, Simeon Peter, because the accused mistook Peter for a member of a rival gang who had robbed him earlier.[145] Totten's evidence was excluded at the first trial in 2007 and Abbey was acquitted.[146] In 2009, the Court of Appeal for Ontario held that

142 Garrett & Neufeld, *supra* note 4 at 34.

143 See *R v Bornyk*, 2013 BCSC 1927 at 39–58 [*Bornyk* Sup Ct 2013].

144 *R v Abbey*, 2007 CarswellOnt 1402 at para 21, 2007 CanLII 1872 (Ont Sup Ct) [*Abbey* Sup Ct 2007]. See also *R v Abbey*, 2017 ONCA 640 at para 41 [*Abbey* CA 2017] (for a concise summary of the evolution of Totten's evidence); Chin, "*Abbey* Road", *supra* note 22 (for a more thorough accounting of *Abbey*).

145 *Abbey* Sup Ct 2007, *supra* note 144 at paras 1–3. See also *R v Abbey*, 2009 ONCA 624 at paras 1–2, 10–11 [*Abbey* CA 2009].

146 *Abbey* CA 2017, *supra* note 144 at paras 13–15.

Totten's evidence should have been admitted and ordered a retrial.[147] Abbey was convicted at that retrial.[148] In 2017, the Court of Appeal ordered a third trial after fresh evidence cast serious doubt on Totten's research program.[149]

The fresh evidence came from an unlikely source: the Crown's own cross-examination of Totten in a subsequent case.[150] That subsequent case was *R v Gager*, where Totten was proffered by the defence to opine that it was unlikely that the two co-accused were gang members.[151] The Crown in *Gager* found several methodological issues that had not been explored in any of the preceding *Abbey* trials but that were deeply relevant to its fair adjudication.[152] *Gager* ultimately provided Abbey with the fodder to bring an appeal based on fresh evidence in 2017.[153] In that decision, Justice Laskin (with Justices Doherty and Roberts concurring) described several weaknesses in Totten's evidence:[154]

- Totten had double-counted some of his interviews across published studies and, thus, he had interviewed fewer individuals than he had originally suggested;[155]
- Totten used a shifting definition of "gang member" to make it seem as if he had interviewed more such individuals than he had;[156]
- Some of the studies Totten relied on did not record whether the individual had been convicted of a homicide, and interview protocols did not include questions about tattoos, raising questions about the validity of conclusions based on those features;[157] and

147 *Abbey* CA 2009, *supra* note 145 at para 4; *Abbey* CA 2017, *supra* note 144 at paras 16, 13–41 (for a summary of the proceedings).

148 *Abbey* CA 2017, *supra* note 144 at paras 16, 36.

149 *Ibid* at paras 138–55.

150 *Ibid* at para 5.

151 See 2012 ONSC 1472 at paras 30–85 [*Gager*].

152 *Ibid* at paras 29–96; *Abbey* CA 2017, *supra* note 144 at para 38.

153 *Abbey* CA 2017, *supra* note 144 at para 45.

154 *Ibid* at paras 69–106.

155 *Ibid* at paras 101–04.

156 *Ibid* at paras 72–82.

157 *Ibid* at paras 66, 86, 92, 98–99 (it may be that these questions were asked or it may be that Totten was, after the fact, remembering them in the way most beneficial to his client's case). See also Nosek, Spies & Motyl, "Utopia", *supra* note 89 at 617 ("Instead, we might remember the gist of what the study was and what we found…Forgetting the details provides an opportunity for reimagining the study purpose and results to recall and understand them in their best (*i.e.* most publishable) light. The reader may, as we do, recall personal examples of such motivated decisions—they are entirely ordinary products of human cognition.").

- None of the above was verifiable because Totten had destroyed his data.[158]

The Court of Appeal admitted the fresh evidence, allowed the appeal, and ordered a third trial.[159] Justice Laskin held that the fresh evidence casting doubt on Totten's opinion was cogent enough that it would have warranted Totten's exclusion if available at trial.[160] More specifically, the evidence undermined the opinion's reliability and, thus, its probative value.[161] In May 2018, Abbey pled guilty to manslaughter and was sentenced to time served, plus one day.[162]

The experience with Mark Totten's evidence in *Abbey* and *Gager* demonstrates how easy it is to mislead the factfinder when using pre-open scientific (yet, in some cases, generally accepted) practices.[163] First, open science would have made it easier for the defence to identify the weaknesses in Totten's analysis. Second, it may have encouraged him to conduct his research more transparently to begin with. Indeed, as three researchers on the vanguard of open science have noted, the mere thought of showing how the sausage is made can inspire vigilance: "Public data posting not only allows others to verify the accuracy of the analyses, but also incentivizes authors to more carefully avoid errors."[164] Adherence to open scientific methods may have also made Totten's evidence more credible, providing valid explanations for what the Court interpreted as self-serving changes to his methodology.

---

158 *Abbey* CA 2017, *supra* note 144 at para 95. But see *ibid* at para 96 (this issue arose in the *Gager* trial, but the Court of Appeal held that it was relevant to Abbey's fresh evidence application); *ibid* at paras 105–06 (in addition, note that Totten's methodology was also unclear in that he claimed to have attended or conducted all interviews but there was not enough time for this to occur).

159 *Ibid* at para 155.

160 *Ibid* at para 109.

161 *Ibid* at paras 117–25.

162 See Betsy Powell, "Accused in Teardrop-Tattoo Case Pleads Guilty to Manslaughter, Released After Almost 11 Years", *The Star* (28 May 2018), online: <www.thestar.com/news/gta/2018/05/28/accused-in-teardrop-tattoo-case-pleads-guilty-to-manslaughter-released-after-almost-11-years.html>.

163 One might also argue that *Abbey* demonstrates the problems that result from conflating exploratory fieldwork with confirmatory (and inculpatory) evidence. At the very least, there should have been some mechanism in place to explain this difference to the factfinder.

164 Nelson, Simmons & Simonsohn, "Psychology's Renaissance", *supra* note 2 at 525. See Jelte M Wicherts & Marjan Bakker, "Publish (Your Data) Or (Let the Data) Perish! Why Not Publish Your Data Too?" (2012) 40:2 Intelligence 73 at 74.

Open data and preregistration of methodology may have made it easier to spot differences between Totten's original research and what he presented in court. For instance, it would have made it more apparent that Totten was double-counting participants (*i.e.* treating the same "gang member" as two data points). Specifically, if each study was recorded on the OSF with participants receiving anonymous identifiers, it would be easier to see the overlap in those identifiers. Note that with sensitive research like Totten's, investigators will have to carefully navigate ethical and confidentiality concerns when considering the openness of their data. For instance, simply using anonymous identifiers may not fully protect those in vulnerable populations. But investigators may consider sharing narratives stripped of identifying information or sharing some data to third-party protected repositories.[165] Indeed, the NASEM recently stated that confidentiality concerns will be one of open science's most significant hurdles—but work is already underway to address these concerns.[166] Importantly, investigators should not simply assume that any level of open data is untenable.

Similarly, preregistration of Totten's definition of "gang-member" (or a list of possible definitions) would have helped. Recall that Totten drew rebuke for changing his definition of "gang member" over time to suit his conclusions. The trial judge in *Gager* also found that an additional category of gang-affiliation proposed by Totten in that case suggested bias.[167] For the defence, Totten had suggested some individuals (implying that this characterization applied to the accused) were not gang members, but "long term friends" of gang members.[168] The trial judge was not impressed with this seemingly *ad hoc* formulation: "The reason Dr. Totten's postulation of the 'long-term friend' suggests bias is that the witness indicated in his cross-examination that he began contemplating the 'long term-friend' category of relationship at the beginning of his doctoral research in the 1990s, but had never committed this idea to paper prior to preparing his report in this case."[169] Preregistering his conceptualization (prior to initial data collection) of what it means to be in a gang (or to be a long-term friend) would

165 See e.g. David Mellor, "Approved Protected Access Repositories" (14 November 2018), online: *OSF Home* <osf.io/tvyxz/wiki/8.%20Approved%20Protected%20Access%20Repositories>.

166 NASEM, "Open Science Report", *supra* note 1 at 50–53, 83–84.

167 *Gager*, *supra* note 151 at para 74.

168 *Ibid*.

169 *Ibid*.

have clarified Totten's evidence a great deal and could have bolstered its credibility.[170]

Finally, the defence and jury would want to know what was inside Totten's file drawer. As we discussed above, the sciences have long suffered from a bias in publication known as the "file drawer effect."[171] This term describes the fact that interesting findings are published, while others languish in the researcher's archives (now, likely a hard drive). Without open scientific reforms, there is simply no way to know how many interviews Totten conducted that did not make their way into his expert report. Accordingly, there is no way to know if some of those interviews were with individuals who got a tattoo for a reason other than killing a rival gang-member.

In parallel with *Abbey*, expert forensic evidence in the recent *Bornyk* proceedings in British Columbia revealed similar concerns in an area far afield from sociology.[172] The expert in *Bornyk*, a fingerprint examiner, identified the accused as the source of a print found at the crime scene.[173] In doing so, he cast fingerprint analysis as a rigorous and objective procedure that admitted of no error: "There's no errors allowed in fingerprint identification. That continues today. There's no errors permitted in fingerprint identification."[174] He did not mention the NRC Report, then four years old, nor the many studies finding that well-trained fingerprint examiners are susceptible to motivated reasoning.[175] He also failed to disclose weaknesses in his own comparison.[176]

Like the Crown's surprising cross-examination of Totten in *Gager*, the *Bornyk* trial also deviated from the usual path. The trial judge apprised himself of several leading reports about fingerprint identification, including the NRC Report.[177] The trial judge subsequently concluded that a "number of troubling aspects" arose from the expert's initial testimony.[178]

---

170 Nosek et al, "Preregistration", *supra note* 44 at 2602 (and preregistration can also assist when the data has already been collected).

171 See generally Dickersin, *supra* note 93.

172 See *Bornyk* Sup Ct 2013, *supra* note 143. See also Gary Edmond, David Hamer & Emma Cunliffe, "A Little Ignorance Is a Dangerous Thing: Engaging with Exogenous Knowledge Not Adduced by the Parties" (2016) 25:3 Griffith L Rev 383 at 387–93 (for a full review of the *Bornyk* case) [Edmond, Hamer & Cunliffe, "Exogenous"].

173 *Bornyk* Sup Ct 2013, *supra* note 143 at para 18.

174 *Ibid* at para 23.

175 *Ibid* at paras 17–32.

176 *Ibid* at paras 55–58; Edmond, Hamer & Cunliffe, "Exogenous", *supra* note 172 at 393–97.

177 *Bornyk* Sup Ct 2013, *supra* note 143 at paras 32–33.

178 *Ibid* at para 39.

These aspects included the expert failing to disclose the subjectivity of fingerprint analysis, the role of unconscious bias, "unexplained discrepancies" between the latent (*i.e.* found) print and Bornyk's print, and discrepancies between prints taken from Bornyk at various times.[179] The trial judge concluded that he was not convinced beyond a reasonable doubt and acquitted Bornyk.[180] The Court of Appeal then reversed that judgment because the trial judge should not have taken judicial notice of the authoritative fingerprint reports.[181] Bornyk was convicted at the retrial.[182]

Abbey and *Bornyk*, cases drawing expertise from seemingly very different fields, both demonstrate the importance of openness through "before and after" comparisons. In *Abbey*'s "before," traditional (closed), scientific practices portrayed Totten's research as highly probative of the meaning of teardrop tattoos. Under the surface, however, lurked uncertainties that the factfinder could not know. But once Totten's practices were exposed (the "after"), it became clear he had taken advantage of the undisclosed flexibility to frame his data in a misleading way. This resulted in the outright exclusion of his evidence and a third trial.

Bornyk is quite similar. There, the "before" constitutes the evidence prior to the trial judge becoming aware of the authoritative reviews of fingerprint identification. At this point, the evidence seemed extremely convincing: the analysis was objective, examiners were error-free, and motivated reasoning did not happen. The "after," as described above, was quite different and seemed to substantially decrease the weight the trial judge accorded the evidence (although the second trial judge, also aware of these limitations, found it convincing enough to convict).[183]

It is important to note that academics have often urged the forensic sciences to behave more like the other sciences. For instance, a group of scientists, academics, and forensic practitioners provided the following prescription: "The simplest advice we can offer to forensic practitioners is to use mainstream scientific methods and norms."[184] But here we saw a forensic scientist comporting himself in a very similar manner to Totten, and to many other mainstream scientists. We think it is time to amend

179 *Ibid* at paras 39–58.

180 *Ibid* at para 161.

181 See *R v Bornyk*, 2015 BCCA 28 at paras 10–11, 20.

182 See *R v Bornyk*, 2017 BCSC 849 at para 144.

183 *Ibid*.

184 Edmond et al, "Model Forensic Science", *supra* note 16 at 497. See also Saks & Faigman, *supra* note 3 at 150 (this entreaty was also implicit in Saks and Faigman's label "the non-science forensic sciences").

the conventional wisdom: expert witnesses should aim not just to comport themselves as scientists, but rather *the most open, transparent, and methodologically rigorous scientists*.

B. Open Science and the Duty to Fairly Present the Case: The Role of the Prosecution

The responsibility for ensuring that the expert's evidence is presented fairly does not end with the expert. This is especially true in criminal matters in which the party tendering the expert is the Crown.[185] In this subsection, we will suggest that the Crown's disclosure obligation should be construed very broadly and that it should take special care when tendering scientific evidence. This is because in science, it has proven much easier than most thought to present strong evidence of weak or non-existent effects by failing to disclose important limitations of the research. Similarly, in law, many wrongful convictions are attributable to the prosecution's failure to disclose exculpatory evidence that weakens its case.[186] In other words, these were failures of transparency, and, thus, open science reforms aimed at improving transparency may be helpful. In this vein, we will conclude this subsection with concrete reforms modeled off those that scientific journals are using to nudge researchers towards more full disclosure.

Most fundamentally, lessons from the open science movement indicate that the scope of the Crown's disclosure obligation should be wider than it has been construed in recent jurisprudence. The duty to disclose all relevant evidence was recognized, in broad language, in the Supreme Court's landmark decision in *R v Stinchcombe*.[187] That duty, however, has been unevenly applied to the Crown's scientific evidence (especially when it is held by the police). For example, in *R v Taillefer*, the Supreme Court

185 See *R v Boucher*, [1955] SCR 16 at 24, 110 CCC 263; *R v Stinchcombe*, [1991] 3 SCR 326 at 333, 68 CCC (3d) 1 [*Stinchcombe*]; For a general review of the risks that flow from unchallenged scientific evidence tendered by the prosecution, see Keith A Findley, "Innocents at Risk: Adversary Imbalance, Forensic Science, and the Search for Truth" (2008) 38:3 Seton Hall L Rev 893.

186 See e.g. Donald Marshall Inquiry, *supra* note 97; Saskatchewan, "Report of the Commission of Inquiry into the Wrongful Conviction of David Milgaard (Hon Mr Justice Edward P MacCallum)", online (pdf): *Government of Saskatchewan* <www.publications.gov.sk.ca/freelaw/Publications_Centre/Justice/Milgaard/Milgaard.pdf>; Manitoba, "The Inquiry Regarding Thomas Sophonow: The Investigation, Prosecution, and Consideration of Entitlement to Compensation" (2001), online (pdf): *Manitoba Justice* <digitalcollection.gov.mb.ca/awweb/pdfopener?smd=1&did=12713&md=1>; *Phillion*, *supra* note 96.

187 *Stinchcombe*, *supra* note 185 at 336.

corrected a post-*Stinchcombe* appellate decision that had found that the Crown's failure to disclose a forensic dentist's earlier opinion, linking the crime to an individual that was not the accused, had not impacted trial fairness.[188] The Supreme Court quashed the conviction and the accused eventually recovered in a civil action against the Attorney General.[189] Recently, however, the Supreme Court held in *R v Gubbins* that maintenance records of breathalyzers do not fall within the Crown's disclosure obligation, despite some evidence that they are relevant to their current operation.[190]

Gubbins is problematic because it runs counter to the meta-scientific discoveries we reviewed above. The records sought bear on the reliability of breathalyzer evidence, and motivated reasoning makes it easy for examiners to dismiss previous errors and mistakes as happenstance and, thus, fail to proactively disclose them.[191] In other words, courts are providing too much discretion to the experts. In such cases, open science would limit that discretion and prescribe transparency: broad disclosure of the entire foundation (or lack thereof) of the scientific case against the accused. There is some precedent for this in the case law. One trial court noted: "The expert's report and any materials which contributed to the foundation of the report or which are clearly relevant to the witness's credibility must be disclosed."[192] Furthermore, while the *Gubbins* majority worried about the inefficiency of ordering disclosure in all cases, at least one crime laboratory in the US has found that "radical transparency" actually improved its efficiency.[193] That lab implemented an automatic system for making lab results available to public defenders.[194] *Gubbins*

188 *Taillefer*, *supra* note 97.

189 *Duguay c Québec (PG)*, 2013 QCCS 4120.

190 See 2018 SCC 44 at paras 2, 45 [*Gubbins*].

191 *Ibid* at para 79. See Rebecca Wexler, "Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System" (2018) 70:5 Stan L Rev at 1343 at 1393–94 (similarly, early challenges to breathalyzers were thwarted by trade secret privilege—courts that denied the privilege ultimately found that the tests were bugged).

192 See *R v Friskie*, [2001] SJ No 216 at para 27, 49 WCB (2d) 375. See *R v Stone*, [1999] 2 SCR 290 at para 99, 134 CCC (3d) 353 (And, on the issue of privilege, Justice Binnie (dissenting, but agreeing with the majority on this point) said, "once a witness takes the stand, he/she can no longer be characterized as offering private advice to a party. They are offering an opinion for the assistance of the court. As such, the opposing party must be given access to the foundation of such opinions to test them adequately.").

193 *Gubbins*, *supra* note 190 at para 53. See also Nicole Bremner Cásarez & Sandra Guerra Thompson, "Three Transformative Ideals to Build a Better Crime Lab" (2018) 34:4 Ga St U L Rev 1007 at 1046.

194 Cásarez & Thompson, *supra* note 193 at 1045–46 (the ultimate goal of this program is a password-protected online portal for public defenders).

removes the motivation for police to make similar advances that may have long-term benefits.

The open science movement also suggests narrowing any litigation privilege[195] that may exist between the Crown and expert as to the foundations of the scientific opinion. No privilege is absolute and litigation privilege may be overridden by competing interests.[196] The accused's interest in making full answers and defences in light of known limitations in scientific reporting may qualify as such a competing interest. Somewhat analogously, Jacob Sherkow, in the US patent context, has cogently argued that the lessons from the open science movement suggest that courts should broaden their view of what is admissible evidence in patent disputes.[197] In short, Sherkow argues that evidence accumulated after the patent was awarded tending to show the initial findings were not reproducible is relevant and should be admitted. Under the current rules, such evidence is typically not admitted.

A goal of any reform should be to remove as much subjective judgment from the disclosure decision as possible. For example, Justice Goudge, observing many cases in which the Crown counsel was silent despite several warning signs in the conduct of Charles Smith, highlighted the need for "detailed guidelines or protocols" to help determine when "issues with expert witnesses" should be disclosed.[198]

Some reforms being implemented in science provide ideas for the content of the guidelines Justice Goudge proposed. These scientific developments require, or incentivize, scientists to positively confirm that they have disclosed their data and not changed their procedure in any way.[199] The reason such acknowledgments work is because it is psychologically more difficult to actively mislead than it is to simply remain quiet about the weaknesses of one's research.[200] The Crown may wish to use similar questions with its experts to ensure that it is receiving an accurate

195 See *Lizotte v Aviva Insurance Company of Canada*, 2016 SCC 52 at para 19 (litigation privilege is the immunity from disclosing documents and communications when their dominant purpose is the litigation at hand).

196 See *General Accident Assurance Co v Chrusz* (1999), 45 OR (3d) 321 at 362–63, 180 DLR (4th) 241 (CA) (and the authorities relied on therein).

197 See Jacob S Sherkow, "Patent Law's Reproducibility Paradox" (2017) 66:4 Duke L J 845.

198 Goudge Report, *supra* note 4 at 454.

199 See Mallory C Kidwell et al, "Badges to Acknowledge Open Practices: A Simple, Low-Cost, Effective Method for Increasing Transparency" (2016) 14:5 PLoS Biology 1 at 3.

200 See Nina Mazar & Dan Ariely, "Dishonesty in Everyday Life and Its Policy Implications" (2006) 25:1 J Public Policy & Marketing 117.

representation of the evidence. Or, in the spirit of preregistration, the Crown may wish to require that the expert produce a pre-determined analysis plan and record any deviations from that plan. Deviations from that pre-specified criteria would be disclosed as a matter of course, and the implications of that change could be openly evaluated in light of possible bias. Similarly, any decision to exclude observations or other data should be made explicit in the expert report.

These enhanced protocols should also be applied to the experts themselves. In this regard, the rules of court across Canada diverge significantly. For example, the *Federal Courts Rules* demand that experts, in their reports, provide "any caveats or qualifications necessary to render the report complete and accurate, including those relating to any insufficiency of data or research and an indication of any matters that fall outside the expert's field of expertise."[201] The rules in Ontario, however, are less insistent and specific.[202] In the "Motherisk Report,"[203] Justice Beaman noted the difference between the more searching federal rules and laxer *Family Law Rules*[204] and suggested the divergence contributed to the miscarriages of justice she studied: "The Rules do not require experts to include any information about the scientific limits of the method they are using, the possibility of contamination, or other issues that could affect the reliability of the opinions or test results. Had these requirements been in place, lawyers and judges may have been alerted to the need to probe the reliability of the Motherisk testing."[205] Like Justice Beaman, we suggest that rules modeled after the *Federal Court Rules* should be adopted across Canada. In the meantime, similar disclosure and acknowledgment that no conflicting

201 See SOR/98-106, Schedule: Code of Conduct for Expert Witnesses, r 52.2, s 3(j) [*Federal Court Rules*]. See also Forensic Evidence Working Group, Council of Judges (Austl), *County Court of Victoria Practice Note: Expert Evidence in Criminal Trials*, No PNCR 1-2014 (24 June 2014) (Hon Elizabeth Gaynor), online (pdf): *County Court of Victoria* <www.countycourt.vic.gov.au/sites/default/files/forms/Practice%20Note%20-%20Expert%20Evidence%20in%20Criminal%20Trials_FINAL%20%28June%202014%29_0.pdf> (similar rules exist in Victoria, Australia); *Uniform Civil Procedure Rules 2005* (NSW), Reg 418, Schedule 7: Expert Witness Code of Conduct.

202 See *Rules of Civil Procedure*, RRO 1990, Reg 194, r 4.1.01(1) (under *Courts of Justice Act*, RSO 1990, c C-43) ("It is the duty of every expert engaged by or on behalf of a party to provide evidence in relation to a proceeding under these rules, (a) to provide opinion evidence that is fair, objective and non-partisan; (b) to provide opinion evidence that is related only to matters that are within the expert's area of expertise; and (c) to provide such additional assistance as the court may reasonably require to determine a matter in issue").

203 Beaman Report, *supra* note 57.

204 O Reg 114/99 (under *Courts of Justice Act*, RSO 1990, c C-43).

205 Beaman Report, *supra* note 57 at 109.

results exist could also be turned into a carrot by awarding expert reports "badges." In science, some journal editors use badges to encourage authors to engage in open scientific methods. For instance, authors may earn badges if they make their data or methods open, and if they preregister their research plan and expectations.[206] This badge method is finding considerable success within science, with one study showing a ten-fold increase in open data after a journal instituted an open data badge.[207]

So, how might badges work in law? Mimicking preregistration, a court may wish to give a badge to an expert report when the expert engages in a pretrial meeting in which the judge, expert, and parties agree on the expert's analysis plan. In the context of real estate valuation, such a plan may include stating the scope of comparison properties that will be used to establish value. This way, the expert could not change the scope when he or she realizes the comparison properties are higher or lower in value than expected (and thus, ostensibly, conflicting with the client's position). As in science, experts wishing to ensure that their reports be viewed as fully credible (in an adversarial system) would be highly motivated to earn these badges.

Finally, if experts do not disclose the data and resources that support their opinions, this may call into question the admissibility of their evidence. For instance, police officers who do not disclose the source material for their opinion because of confidential informant privilege or, more generally, because it is hearsay, may see the probative value of their opinion reduced to the degree that it is inadmissible.[208] We will now turn to this question of admissibility and how it may be informed by open science principles.

C. Open Science and Expert Evidence Law: The Role of the Gatekeeper

In the face of evidence that defies rational evaluation, the next line of defence is exclusion or limitation of that evidence. Here, we will focus on two open scientific lessons for judicial gatekeepers faced with that

206 Center for Open Science, "Badges to Acknowledge Open Practices", online: *OSF Home* <osf.io/tvyxz/wiki/home/>.

207 Kidwell et al, *supra* note 199.

208 See *R v Giles*, 2016 BCSC 294 at paras 165–73; *R v O(F)*, 2016 ONSC 724 at paras 14, 17. See also Wexler, *supra* note 191 at 1394, n 281 (in the US, a court came to a similar decision when trade secret privilege prevented the disclosure of the algorithm behind a DNA test).

decision. First, it is perilous to defer too heavily to practices that are generally accepted by the scientific community. Within science, these practices introduced substantial error into the literature that are only now being corrected—many scientific disciplines are still not open.[209] Rather, gatekeepers should look to the best practices (which often coincide with open practices). Second, courts should attempt to limit the biasing effect of eminence information and instead focus on proficiency, when possible.

Despite Canadian courts never endorsing a general acceptance standard (*i.e.* presuming evidence is reliable if it is accepted in the expert community from which it came),[210] flexibility in the current rules allows considerable deference to generally accepted science. This deference has been expressed in two ways.

First, recall that under *White Burgess*, trial judges are directed to give additional reliability scrutiny to expert evidence when it is "based on novel or contested science or science used for a novel purpose."[211] Therefore, one method of avoiding this scrutiny (and applying a superficial sheen of reliability) is to portray the expertise as old hat.[212] However, as Peter Sankoff has noted, the novelty criterion itself is often unhelpful because new methods and findings (*e.g.* the open science movement) can cast doubt on established findings.[213] As a result, "contested" science seems a better trigger for scrutiny. Troublingly, Chief Justice McLachlin omitted the contested language in her recent enunciation of the expert evidence rules in *R v Bingley*.[214]

Second, courts sometimes construe social scientific evidence and some forensic evidence as "specialized knowledge" and then generally defer to the expert.[215] This move tracks back to the first *Abbey* appeal, which was a remarkably influential decision. The Court found that Totten's method was not scientific, but rather "specialized knowledge gained through

209 NASEM, "Open Science Report", *supra* note 1 at 2.

210 See *R v Singh*, 1993 CarswellBC 3097 at 15, 23 WCB (2d) 558 (BC Sup Ct) (even pre-*Mohan* decisions refused to adopt the *Frye* test: "The American *Frye* test [citation omitted], for the admissibility of new scientific evidence, general acceptance in the scientific community, is now seldom used in the United States and it is not the law in Canada").

211 *White Burgess*, *supra* note 71 at para 23.

212 See Jason M Chin & Helena Likwornik, "*R v Bingley* and the Importance of Scientifically Guided Legal Analysis" (2017) 43:1 Queen's LJ 33 at 49–50 (arguing that there has never been a consistent definition of "novel science," so it is easy to fit a wide spectrum of evidence into that definition).

213 Mewett & Sankoff, *supra* note 19 at 16.5(c)(i).

214 See 2017 SCC 12 at para 17 [*Bingley*]; Chin & Likwornik, *supra* note 212.

215 Chin, "*Abbey* Road", *supra* note 22.

extensive research."[216] As a result, the *Daubert* test was inapplicable: "It was not scientific. It was not novel. And it was not a theory."[217] This method of sidestepping the expert evidence rules has been used to admit a variety of dubious evidence.[218] As we will discuss below, open research, whether or not the subject matter of that research is something courts will characterize as science, is often more credible than that which is conducted opaquely. And, any lack of credibility in that evidence will be easier to spot.

The 2009 *Abbey* decision did not actually hold that "specialized knowledge" should avoid all scrutiny. Rather, Justice Doherty provided nine questions relevant to the reliability of such evidence.[219] In light of the meta-scientific research scaffolding the open science movement, some of these questions are more helpful than others. The questions ranged from field-related factors (*e.g.* whether it is an accepted field and whether the examiner, using accepted methodologies, is accepted in that field) to those focused on the transparency of the methods and data themselves (*e.g.* whether the data is available for scrutiny and whether the methods are susceptible to critical evaluation). After listing these questions, Justice Doherty justified them on the basis of a quote from *Kumho*:

> The objective of that requirement [the gatekeeper function] is to ensure the reliability and relevancy of expert testimony. *It is to make certain that an expert, whether basing testimony upon professional studies or personal experience, employs in the courtroom the same level of intellectual rigour that characterizes the practice of an expert in the relevant field.* [220]

A focus on that quotation from *Kumho* is problematic. It is the part of the judgment that has drawn the most criticism. Influential academics have noted that American decisions are often unduly focused on this quote, characterizing such judgments as, "a dangerous trend and one certainly not endorsed by the *Kumho Tire* Court."[221]

Indeed, a review of those nine questions finds hints of general acceptance in several of them. There are, for instance, references to accepted methodologies and honouring the boundaries of the discipline. The

216 *Abbey* CA 2009, *supra* note 145 at para 108.

217 *Ibid* at para 116.

218 See generally Chin, "*Abbey* Road", *supra* note 22 at 441–48.

219 *Abbey* CA 2009, *supra* note 145 at para 119.

220 *Ibid* at para 120, citing *Kumho*, *supra* note 54 at 152 (the emphasis and British-English spelling of rigour was added by the *Abbey* Court).

221 See David L Faigman et al, *Modern Scientific Evidence: The Law and Science of Expert Testimony*, 2018–19 ed (Thomson Reuters, 2018) at §1:28.

answers to such questions may be misleading. Recall, for instance, that the journal, *Nature*, expressly admitted that it had not historically published enough information for peer reviewers to do their jobs.[222] Similarly, we have reviewed many flexibilities in the research process that allow researchers to frame their findings to seem more credible than they are. These were then, and in many cases, still are, generally accepted practices.[223] Instead, experts in court should be held to the best research practices in their field, which are often open scientific standards.[224]

Similarly, the application of the question, "[t]o what extent is the proffered opinion based on data and other information gathered independently of the specific case or, more broadly, the litigation process?",[225] should be informed by open science research. As we noted in Part I, courts and academics regularly draw this distinction between research that is conducted for the purposes of litigation and research that is not. They suggest that litigation-driven research is more likely to be biased.[226] In this vein, Justice Doherty was, to some extent, correct in implying that evidence collected in the course of a criminal investigation is often subject to cognitive biases.[227] The same, however, is true of data collected long before litigation is contemplated. For instance, meta-scientists have documented pervasive use of the researcher degrees of freedom on topics far afield from litigation.[228] This was demonstrated by Totten's research in *Abbey* itself—his data were collected long before the *Abbey* case.

---

222 Reducing Our Irreproducibility, *supra* note 39 at 398.

223 NASEM, "Open Science Report", *supra* note 1 at 2, 12 ("Sharing data, code, and other research products is becoming more common, but is still not routinely done across all disciplines").

224 Chin, "Replicability Crisis", *supra* note 12 at 225; *Abbey* CA 2017, *supra* note 144 at para 124 (note that not all of Totten's practices should be classified as generally accepted; the 2017 decision generally characterized him as dishonest, but expressly refrained from making any finding of research fraud).

225 *Abbey* CA 2009, *supra* note 145 at para 119.

226 Haack, *supra* note 22. See also *Daubert v Merrell Dow Pharmaceuticals, Inc*, 43 F (3d) 1311 at para 24 (9th Cir 1995).

227 See Gary Edmond & Mehera San Roque, "Quasi-Justice: Ad Hoc Expertise and Identification Evidence" (2009) 33 Crim L J 8 at 32–33. See also *Gillham v R*, [2012] NSWCCA 131 at paras 158–98 (the miscarriage of justice in the case of Jeffrey Gilham in which the trial court admitted unrepresentative "experiments" performed by the Crown's expert to demonstrate the rate at which fire spreads).

228 John, Loewenstein & Prelec, *supra* note 36. See also Kellia Chiu, Quinn Grundy & Lisa Bero, "'Spin' in Published Biomedical Literature: A Methodological Systematic Review" (2017) 15:9 PLoS Biology 1 (research reports can also "spin" results to fit with the desires of the author).

Indeed, Susan Haack has suggested that the litigation-driven characterization, while potentially useful, is a messy one. In short, she argues that there is greater potential bias in litigation-driven research, but that does not mean that the evidence is unhelpful or invalid in many cases.[229]

Open scientific research lends additional force to Haack's comment and suggests that transparency presents a way to distinguish the credible litigation-driven research from the more dubious sort. For instance, there is a concern that drug company-funded research examining the safety of a drug will present a misleading picture of the science because the company indirectly controls the data collection, analysis, and peer-review process.[230] Open scientific standards, while certainly not a panacea for this bias, would assist in assessing the credibility of research; findings and conditions that did not support the company's preferred outcome would be more difficult to suppress.[231] Indeed, vocal critics of the pharmaceutical industry have noted that large analyses of the medical literature (like that performed by the Cochrane Collaboration) are difficult to perform when research is systematically suppressed (*i.e.* preclinical research that is not preregistered).[232]

Consider another distinction, between qualitative and quantitative research, made in the *Abbey* judgments. In the 2009 appeal, Justice Doherty appeared to accept Totten's insistence that concepts from *Daubert*, like error rates and testing, were inapplicable to qualitative research methods, like his interviews.[233] However, at the 2017 appeal, Justice Laskin revisited this distinction, noting that the qualitative and quantitative parts of Totten's research were intertwined.[234] Totten's conclusions drew force from the number of interviews he performed.

Justice Laskin's insight deserves further discussion. Much of the research we have relied on in our review of the open science movement

229 Haack, *supra* note 22 at 1077.

230 *Ibid* at 1067–69.

231 And, preregistration would help distinguish between pre- and post-diction. For instance, consider a study that found a drug had no overall effect on birth defects. However, if the cohort was divided up based on when the mother took the drug, there was a finding that at certain times, the drug did seem to increase the chances of birth defects. If the researchers had preregistered the prediction that there were certain sensitive times in which the drug was dangerous, that would significantly increase the credibility of the claim that the drug was dangerous. Otherwise, critics could (persuasively) argue that there is always a way to carve up the data such that there is a period in which it is not safe to take the drug.

232 See e.g. Ben Goldacre, *Bad Science* (London: Fourth Estate, 2008) at 158–59.

233 *Abbey* CA 2009, *supra* note 145 at paras 47, 107–12.

234 *Abbey* CA 2017, *supra* note 144 at para 122.

has been quantitative—the "researcher degrees of freedom" that cause a field's actual numerical error rate to diverge from its reported error rate. But the same logic holds for non-quantitative research because statistics are often just a method of standardized inference.[235] Open and transparent research is both numerically and logically more credible and trustworthy. In other words—and like the above discussion about litigation versus non-litigation driven science—open scientific norms and practices can be applied to both quantitative and qualitative research.

With Totten's research, for example, imagine if he had pre-recorded his definition of "gang member," provided evidence that there were not additional interviews sitting in his file drawer (through preregistration), and pre-recorded his reasons for performing his interviews. To any observer (familiar with inferential statistics or not), his findings would be more credible and better lend themselves to rational evaluation.

Some of the questions provided in *Abbey* address open scientific principles more directly.[236] These questions include whether the expert's reasoning and methods were "clearly explained by the witness and susceptible to critical examination," and whether the data was "accurately recorded, stored, and available."[237] Similarly, meta-scientific research suggests that expert conclusions are more reliable when the authors make the underlying data available and thus open for scrutiny by peers.[238] And methods are indeed better "susceptible to critical examination" when they are fixed before the data are collected (*i.e.* preregistered).

We would, however, rethink the "same intellectual rigour" touchstone and provide a new one to use as an interpretative aid for all nine questions. All expert evidence, whether it originates from hard science, social science, or forensic science, *should be transparently produced and reported*. This touchstone is not a far leap from the two questions we have endorsed, and it should be easy to read the remaining questions with the theme of openness (*e.g.* "accepted and open methodologies"). We do not think that this interpretation should be controversial: openness flows from established principles of expert evidence law, which demand that opinions be

235 See Robert P Abelson, *Statistics As Principled Argument* (New York: Psychology Press, 1995).

236 *Abbey* CA 2017, *supra* note 144 at paras 93, 116.

237 *Abbey* CA 2009, *supra* note 145 at para 119; *Abbey* CA 2017, *supra* note 144 at para 93.

238 See Jelte M Wicherts, Marjan Bakker & Dylan Molenaar, "Willingness to Share Research Data Is Related to the Strength of the Evidence and the Quality of Reporting of Statistical Results" (2011) 6:11 PLoS ONE 1. Similarly, experts in courts have been chided for not sharing their data: *JLJ*, *supra* note 75 at para 57.

presented in way that the factfinder can understand,[239] and the academic work in this field.[240]

If experts are more transparent and gatekeepers are more attuned to open science principles, it will also make the job of cross-examining experts easier. Although we focused in Part V(b) on the party proffering the expert, the cross-examiner also has an important role to play. Excellent work providing tangible recommendations for cross-examining potentially unreliable and biased expertise already exists.[241] Therefore, we simply add that cross-examiners should take guidance from the open science practices described above: did they change their methodology after seeing partial results? Did the peer reviewers (of the scientific foundations of the opinion) have access to their data and detailed research materials? Is their methodology and data stored in a manner that demonstrates it was not changed after beginning the research?

Finally, lessons from open science also suggest that courts should regulate eminence rather than enable it. Eminence is, at best, a noisy measure of any fact relevant to adjudication.[242] Its biasing effect contributes to the replicability crisis and courts should seek to limit its effect.[243] *Gager* provides an example of a (failed) appeal to eminence. In *Gager*, Totten declared himself "an expert witness and Canadian expert on gangs."[244] The trial judge chided him for this, saying "the statement suggests to me that Dr. Totten was characterizing himself as the pre-eminent expert in the field."[245]

A curative may also flow from *Abbey*. In the original *Abbey* appeal, Justice Doherty suggested that the gatekeeper play an active role in establishing the scope of expert evidence: "The trial judge may admit part of the proffered testimony, modify the nature or scope of the proposed opinion, or edit the language used to frame that opinion."[246] This recommendation has been endorsed by the Supreme Court at least twice and it offers an

239 *Mohan*, *supra* note 17 at 21–22; *Béland*, *supra* note 17 at 415–16.

240 Edmond, *supra* note 19; Paciocco, "Expert Testimony", *supra* note 51 at 156.

241 See Gary Edmond et al, "How to Cross-Examine Forensic Scientists: A Guide for Lawyers" (2014) 39 Austl Bar Rev 174; Paciocco, "Jukebox", *supra* note 94 at 599–608.

242 Garrett & Mitchell, *supra* note 134.

243 *Ibid*.

244 *Gager*, *supra* note 151 at para 33.

245 *Ibid* at para 38.

246 *Abbey* CA 2009, *supra* note 145 at para 63. See also *R v Ranger* (2003), 67 OR (3d) 1 at para 63, 178 CCC (3d) 375 (Ont CA).

economical solution to managing eminence.[247] This is because, while it is difficult to ascertain all of the scientific issues that may arise as the expert gives evidence and is cross-examined, eminence is low-hanging fruit for early scoping. In forensic science, for instance, an examiner's proficiency (*i.e.* his or her error rate on similar tasks) is much more important than his or her eminence.[248] For instance, awards and commendations may not necessarily capture the most important variable—how often the examiner tends to be wrong. In such cases, trial judges may find it useful to establish the boundaries of reportable eminence (*i.e.* sticking to proficiency) very early on, such as after a *voir dire* or even during case management meetings.

VI. CONCLUSION: IMPROVING TRUST AND EFFICIENCY IN EXPERT EVIDENCE

In this article, we began by describing the state of openness and transparency in science, our culture's prevailing means of producing knowledge. Meta-scientific research suggests that reforms aimed at openness can help generate a more reliable body of knowledge by curbing researchers' flexibility to come to a more publishable conclusion. We then turned to the legal system, where expert witnesses are generally in charge of generating and conveying knowledge. These experts, some who come from an academic scientific background and some who do not, labour under many of the same demands as mainstream scientists and make similar errors in similar ways. As a result, we suggested that openness and transparency should be treated as a guiding principle for expert evidence, the parties proffering it, and the courts evaluating it.

While we have mainly focused on openness and transparency in expert evidence as a way to improve the reliability of that evidence (or to make it easier to see that it is not reliable), we will conclude with two additional benefits to open and transparent expert evidence: improved trust and efficiency. We hope these two benefits may encourage even the more reluctant adopters to adopt open and transparent practices.

Beginning with trust, recent (and ongoing) controversies may have damaged the public perception of several fields of expert evidence. As to experts generally, they are often accused of being partisan.[249] And in forensic science, concerns about its reliability appear to be reaching a fever

247 See *R v Sekhon*, 2014 SCC 15 at paras 46–48; *Bingley*, *supra* note 214 at para 17.

248 Garrett & Mitchell, *supra* note 134.

249 Paciocco, "Jukebox", *supra* note 94.

pitch. As has been extensively documented, past controversies called into question the credibility of forensic science.[250] Public sentiment finally seems to be catching on to this information, with the media documenting visceral mistakes made by current practices. This includes evocative, searching documentaries and extensive media coverage.[251] Recent empirical research reinforces the view that the public is beginning to see forensic science as less believable.[252]

The ways in which forensic science's two most closely allied fields (*i.e.* mainstream science and law) have dealt with issues of trust may be instructive. In science, the open science movement is surely very much about the reliability and democratization of knowledge. But it also aims to improve the public's trust in science by laying bare its inner workings.[253] Otherwise, a lemons market could result, with consumers of science so uncertain about the quality of findings that they dismiss most results. Similarly, in many jurisdictions, law seeks to maintain accountability, and, thus, trust in its processes through open courtrooms and published judgments.[254]

Experts (or those who generate research regularly used by experts) may wish to follow suit. For instance, although many journals (including forensic scientific journals) do not require publishing data and preregistering analysis plans, there is nothing stopping researchers from doing so. Such practices may assuage concerns that researchers' studies used questionable research practices. As to the application of foundational research to specific cases, forensic labs may also wish to improve trust in their procedures by making them available to the public. In this vein, forensics

250 Saks & Faigman, *supra* note 3; Dufraimont, *supra* note 3; Edmond & Roach, *supra* note 3; Gold, *supra* note 3; Beaman Report, *supra* note 57; Lang Report, *supra* note 58.

251 See e.g. Center for Integrity in Forensic Sciences, "Reform in Forensic Sciences, Crime Laboratories, and the Courtroom" (2018), online: *Center for Integrity in Forensic Sciences* <cifsjustice.org/#/main>; "Forensic Science: Last Week Tonight with John Oliver (HBO)" (1 October 2017), online (video): *Youtube* <www.youtube.com/watch?v=ScmJvmzDcG0>; "Making a Murderer" (18 December 2015), online (video): *Netflix* <www.netflix.com>; "The Staircase" (2018), online (video): *Netflix* <www.netflix.com>; Jennifer L Mnookin, "The Uncertain Future of Forensic Science" (2018) 147:4 Daedalus, J American Academy Arts & Sciences 99 (Jennifer Mnookin suggests forensic science is at a turning point, with reform or staying the course being equally likely).

252 See Gianni Ribeiro, Jason M Tangen & Blake M McKimmie, "Beliefs About Error Rates and Human Judgment in Forensic Science" (2019) 297 Forensic Science Intl 138.

253 See Simine Vazire, "Quality Uncertainty Erodes Trust in Science" (2017) 3:1 Collabra: Psychology 1.

254 See Emma Cunliffe, "Open Justice: Concepts and Judicial Approaches" (2012) 40:3 Federal L Rev 385 at 388.

laboratories may note the successes of the Houston Forensic Science Center (HFSC), which was created after scandals plagued its predecessor organization.[255] The first "transformative ideal" adopted by the HFSC was transparency.[256] This ideal is expressed through measures like public board meetings, direct access for the public defenders' office to lab results, and an online portal open to the public with information about standard operating procedures, incidents, and the responses to those incidents.[257]

Beyond trust, expert witnesses—forensic scientists in particular—could take note of the efficiency gains open scientists have found in some of their reforms.[258] One of the most daunting tasks facing forensic science is the validation of many of its practices through large-scale studies.[259] The OSF, which includes tools for combining research efforts and sharing data,[260] could assist forensic researchers in conducting these studies across multiple labs.[261] As to the day-to-day work of forensic practice, transparency may also beget efficiency in the long run. Here, again, the HSFC reports that their transparency reforms had the ancillary benefit of making their work more efficient.[262] In their case, providing a web portal for public defenders and the public significantly freed up administrative resources and cut down on freedom of information requests: "While the HFSC has pursued radical transparency as a way to strengthen public trust in its operations, its commitment to transparency has resulted in an added benefit: the creation of a more efficient criminal justice system that saves time and money for all participants."[263]

Overall, we have endeavoured to outline the many benefits that may accrue to the legal system if it takes seriously the transparency and openness reforms going on in science. These include efficiency and a system that

255 Cásarez & Thompson, *supra* note 193.

256 *Ibid* at 1042–43.

257 *Ibid* at 1045.

258 Munafò et al, *supra* note 1 at 2.

259 PCAST Report, *supra* note 15 at 52.

260 See e.g. Center for Open Science, "StudySwap: A Platform for Interlab Replication, Collaboration, and Research Resource Exchange" (2011–2019), online: *Center for Open Science* <osf.io/view/StudySwap/>.

261 See Kristy A Martire, Bethany Growns & Danielle J Navarro, "What Do the Experts Know? Calibration, Precision, and the Wisdom of Crowds Among Forensic Handwriting Experts" (2018) 25:6 Psychonomic Bulletin & Rev 2346 (similarly, open scientific methods may assist in developing systems that combine the judgments of multiple examiners and machines—but such work in is in its nascent stages).

262 Cásarez & Thompson, *supra* note 193 at 1046.

263 *Ibid*.

is more trustworthy and accountable. But, perhaps most fundamentally, transparently produced and presented expert evidence helps fulfil the ideals of expert evidence: providing knowledge that is helpful to the court, assailable by the adverse party, and understandable to the trier of fact.