WHERE IS THE EVIDENCE IN EVIDENCE-BASED LAW REFORM?

JASON M CHIN,* MALGORZATA LAGISZ** AND SHINICHI NAKAGAWA**

Law reform bodies frequently express a commitment to evidence-based law and policy recommendations. They also endorse the importance of the transparency and democratisation of their processes. They do not, however, connect these two goals of evidence-based policy and transparency. This stands in contrast to the ongoing revolution in several fields of research towards open science and synthesis, which envisions transparency and openness as a means to improve the reliability of research. We suggest herein that several recent concerns surrounding evidence-based law reform, such as allegations of bias among officials, can be addressed through open science and synthesis. We include a novel study of 22 research syntheses commissioned by the Royal Commission into Institutional Responses to Child Sexual Abuse, finding uneven adoption of even the most basic transparency measures. We end with five proposals that advance transparent evidence-based law reform, including greater disciplinary diversity among law reform staff.

I INTRODUCTION

'*The overarching principle of open science by design is that research conducted openly and transparently leads to better science. Claims are more likely to be credible – or found wanting – when they can be reviewed, critiqued, extended, and reproduced by others.*'[1]

Governments and law reform bodies frequently express a commitment to evidence- and research-based policy and law.[2] They also stress the importance

* School of Law, University of Sydney, Sydney.
** Evolution and Ecology Research Centre and School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney.

1 Committee on Toward an Open Science Enterprise et al, *Open Science by Design: Realizing a Vision for 21st Century Research* (National Academies Press, 2018) 107 (emphasis omitted).

2 E Miguel et al, 'Promoting Transparency in Social Science Research' (2014) 343 *Science* 30, 30; Australian Law Reform Commission, *Annual Report 2019–2020* (2 October 2020) 2 ('*ALRC Annual Report*'); Law Commission for England and Wales, *Annual Report 2019–20* (Report No 391, 14 July 2020) 57 ('*LCEW Annual Report*'); *Royal Commission into Institutional Responses to Child Sexual Abuse* (Final Report, December 2017) vol 1, 42 ('*RCIRCSA Final Report*').

of transparency and public access in their processes.[3] These values, a strong evidence basis and transparency, are often discussed as serving unrelated ends.[4] In this article, we suggest that transparency, besides being valuable for its own sake, is also central to evidence-based policy. In particular, as described in the above epigraph, a 'revolution'[5] over the past decade in several fields of research towards open science[6] and open synthesis[7] represents a promising avenue towards more reliable, democratic, and efficient law reform processes. Open science and synthesis, however, have been almost entirely overlooked by recent law reform efforts and law reform commentators. We seek to remedy that.

The main reason that open science and synthesis (as defined in Part II) are important in the context of law reform is that they can make the evidence gathering and synthesis processes more reliable and less susceptible to bias. More specifically, open and transparent synthesis helps avoid the situation whereby researchers can selectively find and cite (unconsciously or otherwise) the sources and studies that fit their preferred view and present those to law

3 The Australian Law Reform Commission ('ALRC') highlights publication of its reports and public engagement on its website: Australian Law Reform Commission, 'Law Reform Process', *Australian Government* (Web Page) <https://www.alrc.gov.au/about/law-reform-process/>. One part of the website states '[t]he ALRC is committed to publishing information that increases the transparency of our decision-making processes': Australian Law Reform Commission, 'Access to Information', *Australian Government* (Web Page) <https://www.alrc.gov.au/about/access-information/>. Many state commissions in Australia follow a similar pattern of public consultations, publication of a consultation report, accepting further submissions that are often published, and the ultimate publication of a report and much of the underlying material. In the United Kingdom ('UK'), see *LCEW Annual Report* (n 2) 3: 'I also believe our Senior Management Team provides excellent strategic advice to the Board, ensuring that Board decisions are evidence-based, transparent and stand up to wider scrutiny'.

4 Transparency has been described as promoting many salutary ends, but it is difficult to find references connecting it to the reliability of evidence. For instance, transparency has been described as making sure people have a say in the law reform process by giving them access to the source of a commission's recommendations and by making those recommendations understandable: Roderick A Macdonald, 'Law Reform and Its Agencies' (2000) 79(1) *Canadian Bar Review* 99, 104–5; Martin Partington, 'The Relationship between Law Reform and Access to Justice: A Case Study' (2005) 23 *Windsor Yearbook of Access to Justice* 375, 391–2. Similarly, public hearings can serve an educative purpose, raising awareness of the issues being dealt with. 'The Commission also aimed to ensure that it raised awareness about the prevalence, incidence and effects of family violence and about the activities and deliberations of other organisations working to address it. The public hearings were broadcast live over the internet, and submissions, witness statements and hearing transcripts were published on our website': *Royal Commission into Family Violence* (Summary and Recommendations, March 2016) 2. See also Macdonald (n 4) 106; Partington (n 4) 392.

5 See, eg, Joshua D Angrist and Jörn-Steffen Pischke, 'The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics' (2010) 24(2) *Journal of Economic Perspectives* 3 <https://doi.org/10.1257/jep.24.2.3>; Simine Vazire, 'Implications of the Credibility Revolution for Productivity, Creativity, and Progress' (2018) 13(4) *Perspectives on Psychological Science* 411 <https://doi.org/10.1177/1745691617751884>; Editorial, 'Synthesis Revolution' (2020) 4 *Nature Ecology and Evolution* 489 <https://doi.org/10.1038/s41559-020-1180-z>.

6 Committee on Toward an Open Science Enterprise et al (n 1); Marcus R Munafò et al, 'A Manifesto for Reproducible Science' (2017) 1(1) *Nature Human Behaviour* 1, 5 <https://doi.org/10.1038/s41562-016-0021>.

7 See, eg, Neal R Haddaway et al, 'Open Synthesis and the Coronavirus Pandemic in 2020' (2020) 126 *Journal of Clinical Epidemiology* 184 <https://doi.org/10.1016/j.jclinepi.2020.06.032> ('Open Synthesis'); Shinichi Nakagawa et al, 'A New Ecosystem for Evidence Synthesis' (2020) 4 *Nature Ecology and Evolution* 498 <https://doi.org/10.1038/s41559-020-1153-2>.

reform bodies (and can provide similar value as the law reform bodies consider and weigh up the evidence themselves). Indeed, syntheses of evidence have been known to contain biases, even when the work is notionally unconnected to political issues.[8] This is because researchers, like all of us, are susceptible to cognitive biases, such as confirmation bias (ie, focusing on evidence that fits our preferred view) and apophenia (seeing patterns in randomness).[9] In a time when politically motivated attacks on research and science seem higher than ever,[10] the implementations we suggest below may be beneficial in both improving the law reform process itself and maintaining its integrity.[11]

Bringing open and transparent synthesis and science to law reform can also improve its efficiency by reducing waste. Law reform is a costly process, a fact that has drawn criticism and caused several bodies to shutter.[12] One justification for this cost is the knowledge law reform processes produce.[13] For example, in one royal commission we delve more deeply into, the *Royal Commission into Institutional Responses to Child Sexual Abuse* (‘*RCIRCSA*’), the commissioners funded 22 syntheses from external researchers, several original studies, and a large-scale mock jury study.[14] This (publicly funded) knowledge, to fulfil its potential, ought to be verifiable and reusable by other researchers, non-governmental organisations, and future law reform bodies. Here, open synthesis assists: ‘How easy or how

8 See, eg, Matthew J Page, Joanne E McKenzie and Andrew Forbes, ‘Many Scenarios Exist for Selective Inclusion and Reporting of Results in Randomized Trials and Systematic Reviews’ (2013) 66 *Journal of Clinical Epidemiology* 524 <https://doi.org/10.1016/j.jclinepi.2012.10.010>; Matthew J Page et al, ‘Bias Due to Selective Inclusion and Reporting of Outcomes and Analyses in Systematic Reviews of Randomised Trials of Healthcare Interventions’ (2014) 10 *Cochrane Database of Systematic Reviews* MR000035:1–42 <https://doi.org/10.1002/14651858.mr000035.pub2>; Shanil Ebrahim et al, ‘Meta-Analyses with Industry Involvement Are Massively Published and Report No Caveats for Antidepressants’ (2016) 70 *Journal of Clinical Epidemiology* 155 <https://doi.org/10.1016/j.jclinepi.2015.08.021>.

9 Munafò et al (n 6) 1.

10 Allison Orr Larsen, ‘Constitutional Law in an Age of Alternative Facts’ (2018) 93(2) *New York University Law Review* 175; Martin Bush et al, ‘Real Problem, Wrong Solution: Why the Nationals Shouldn’t Politicise the Science Replication Crisis’, *The Conversation* (online, 25 September 2019) <https://theconversation.com/real-problem-wrong-solution-why-the-nationals-shouldnt-politicise-the-science-replication-crisis-124076>.

11 Cary Funk et al, *Trust and Mistrust in Americans’ Views of Scientific Experts* (Pew Research Center Report, 2 August 2019) 24; Simine Vazire, ‘Quality Uncertainty Erodes Trust in Science’ (2017) 3(1) *Collabra* 1 <https://doi.org/10.1525/collabra.74>; Richard C Cottrell, ‘Scientific Integrity and the Market for Lemons’ (2014) 10(1) *Research Ethics* 17 <https://doi.org/10.1177/1747016113494651>.

12 For criticisms, see Michael Mintrom, Deirdre O’Neill and Ruby O’Connor, ‘Royal Commissions and Policy Influence’ (2021) 80(1) *Australian Journal of Public Administration* 80, 81 <https://doi.org/10.1111/1467-8500.12441>; Macdonald (n 4) 100. For shuttered bodies, see Marcus Moore, ‘The Past, Present and Future of Law Reform in Canada’ (2018) 6(2) *Theory and Practice of Legislation* 225, 239–40 <https://doi.org/10.1080/20508840.2018.1476114>.

13 See Mintrom, O’Neill and O’Connor (n 12) 85.

14 ‘Research’, *Royal Commission into Institutional Responses to Child Sexual Abuse* (Web Page) <https://www.childabuseroyalcommission.gov.au/research> (‘RCIRCSA Research Website’). See descriptions of the *Royal Commission into Institutional Responses to Child Sexual Abuse*’s (‘*RCIRCSA*’) research in Mintrom, O’Neill and O’Connor (n 12) 89–91; Jill Hunter and Richard I Kemp, ‘Proposed Changes to the Tendency Rule: A Note of Caution’ (2017) 41 *Criminal Law Journal* 253.

difficult the [synthesis] update is to conduct depends on the transparency of the original'.[15] The same goes for the cases in which law reform bodies pay for data collection (as opposed to synthesis) – the underlying data should be made available for reanalysis and future study. Some have even argued that doing otherwise breaches a social contract between university researchers, who are often publicly funded, and the public. The same argument is even more forceful in the case of law reform funded research, which is wholly publicly funded and collects data of wide societal importance.[16]

From here, Part II will set the stage with a brief introduction to open science and synthesis. Part III then turns to recent attempts at evidence-based policy in law reform and some critiques that have resulted. This includes our study of synthesis reports commissioned by the *RCIRCSA*, which we have coded and analysed to determine whether they meet the standards of open synthesis, and thus made efforts to control for bias and to be updatable as new research is conducted. This review finds a low and uneven level of transparency in the *RCIRCSA*'s syntheses. Part IV suggests five ways in which law reform can be improved through adopting transparent evidence gathering and synthesis, and through greater diversity among law reform staff. Part V concludes with reflections on roadblocks that open law reform may face and suggestions for overcoming those challenges.

Our scope in the following analysis is wide in that we will be discussing law reform efforts generally: both standing bodies like the Australian Law Reform Commission ('ALRC') and ad hoc royal commissions (the latter of which have broader purposes).[17] We recognise the considerable differences between these types of reforms and discuss some of them.[18] That said, much of our analysis and recommendations is broadly applicable. Second, we have not conducted our own systematic, open review of law reform policies and procedures. Rather, we have focused mostly on recent law reform efforts that have attracted commentary elsewhere. This allows our work to build on existing scholarship by suggesting how limitations pointed out by others could be addressed through open science and synthesis methods. Additionally, much of our analysis focuses, out of convenience, on the *RCIRCSA* because of the high (and admirable) degree to

15 Joshua R Polanin, Emily A Hennessy and Sho Tsuji, 'Transparency and Reproducibility of Meta-Analyses in Psychology: A Meta-Review' (2020) 15(4) *Perspectives on Psychological Science* 1026, 1028 <https://doi.org/10.1177/1745691620906416>.

16 Todd J Vision, 'Open Data and the Social Contract of Scientific Publishing' (2010) 60(5) *BioScience* 330 <https://doi.org/10.1525/bio.2010.60.5.2>.

17 See Jacqueline Fuller, 'The David Eastman Case: The Use of Inquiries to Investigate Miscarriages of Justice in Australia' (2020) 45(1) *Alternative Law Journal* 60 <https://doi.org/10.1177/1037969X19886348>; *Royal Commissions Act 1902* (Cth).

18 We also recognise the considerable limitation that governments are not obligated to follow the recommendations of law reform bodies. That said, the recommendations of law reform bodies are often quite influential: Mintrom, O'Neill and O'Connor (n 12) 81.

which it produced research and data.[19] It is perhaps only natural that it fell short in many ways.[20] In other words, while a more thorough review of law reform of Australia and abroad would be worthwhile, the current scope provides a useful lens for analysing the limitations of evidence-based law reform in Australia, and how open science and synthesis may contribute to the discussion. Finally, to assist with some of the technical terms we use in the following section, Figure 1 can be used as a quick reference.

19 The Commission stated at several points its commitment to both evidence and transparency, two values that are central to this article, further reinforcing the utility of focusing some of our efforts on the *RCIRCSA*: *RCIRCSA Final Report* (n 2) 24, 42, 46–7, 54.

20 We also note that we will be reviewing law reform efforts from the last decade to suggest that open and transparent research practices can help future efforts avoid the pitfalls of the past. We discourage assigning blame as to whether those previous efforts were out of touch with best practices at the time they were conducted. For instance, the historical context is complicated, with fields following different timelines with respect to different practices. For example, transparent practices in synthesis were widespread from at least 2012: Tanja Rombey et al, 'A Descriptive Analysis of the Characteristics and the Peer Review Process of Systematic Review Protocols Published in an Open Peer Review Journal from 2012 to 2017' (2019) 19(57) *BMC Medical Research Methodology* 57:1–9 <https://doi.org/10.1186/s12874-019-0698-8>. In psychological primary research (ie, collecting new data) one group suggests that psychology's renaissance (eg, preregistration, larger sample sizes, sample size planning) began in 2010: Leif D Nelson, Joseph Simmons and Uri Simonsohn, 'Psychology's Renaissance' (2018) 69 *Annual Review of Psychology* 511, 512 <https://doi.org/10.1146/annurev-psych-122216-011836>. More generally, transparent practices in primary social science research have grown exponentially in the last decade: Kai Kupferschmidt, 'A Recipe for Rigor' (2018) 361(6408) *Science* 1192 <https://doi.org/10.1126/science.361.6408.1192>; Garret Christensen et al, *Open Science Practices Are on the Rise: The State of Social Science (3S) Survey* (Report, August 2019) <https://doi.org/10.31222/osf.io/5rksu>. And, preregistration of primary research in medicine had a head start, having been required by law for over 15 years: Kay Dickersin and Iain Chalmers, 'Recognizing, Investigating and Dealing with Incomplete and Biased Reporting of Clinical Research: From Francis Bacon to the WHO' (2011) 104(12) *Journal of the Royal Society of Medicine* 532, 534 <https://doi.org/10.1258/jrsm.2011.11k042>; Kay Dickersin and Drummond Rennie, 'The Evolution of Trial Registries and Their Use to Assess the Clinical Trial Enterprise' (2012) 307(17) *Journal of the American Medical Association* 1861 <https://doi.org/10.1001/jama.2012.4230>.

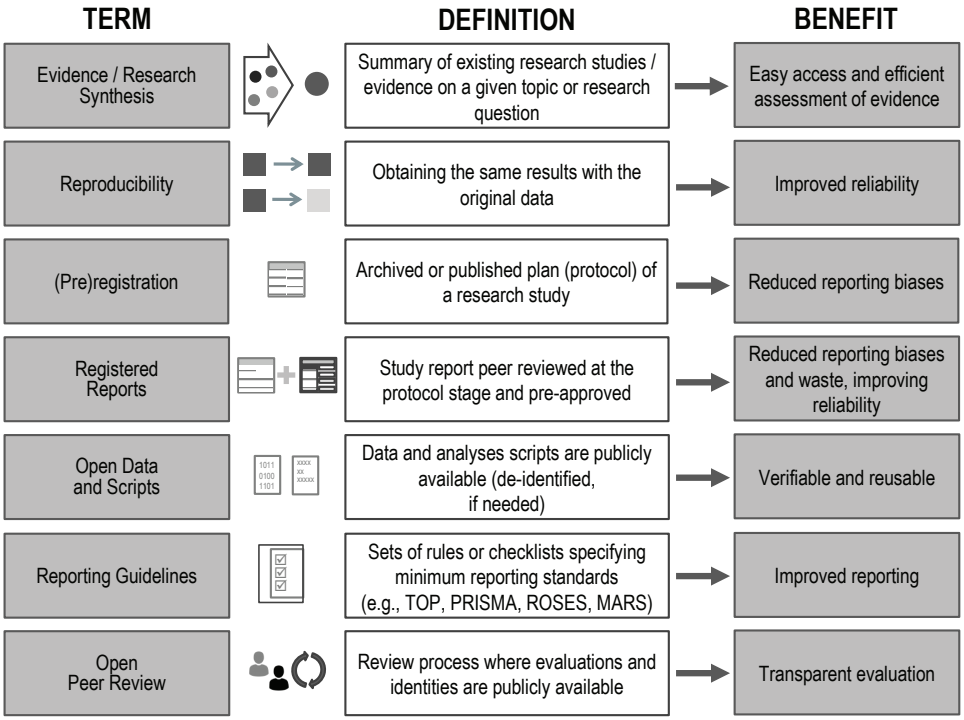

Figure 1: A glossary of key terms used in this article

II METARESEARCH, OPEN SCIENCE, AND OPEN SYNTHESIS

Metaresearch (sometimes called metascience) uses, among other approaches, modern data-driven methods,[21] engagement with scientific communities,[22] and a focus on informing and testing reforms,[23] to study the research process.[24] This field

21 Amanda Kvarven, Eirik Strømland and Magnus Johannesson, 'Comparing Meta-Analyses and Preregistered Multiple-Laboratory Replication Projects' (2019) 4 *Nature Human Behaviour* 423 <https://doi.org/10.1038/s41562-019-0787-z>.

22 Balazs Aczel et al, 'A Consensus-Based Transparency Checklist' 4(1) *Nature Human Behaviour* 4.

23 Mallory C Kidwell et al, 'Badges to Acknowledge Open Practices: A Simple, Low-Cost, Effective Method for Increasing Transparency' (2016) 14(5) *PloS Biology* e1002456:1–15 <https://doi.org/10.1371/journal.pbio.1002456>; Anisa Rowhani-Farid and Adrian G Barnett, 'Badges for Sharing Data and Code at Biostatistics: An Observational Study' (2018) 7 *F1000Research* 90:1–17 <https://doi.org/10.12688/f1000research.13477.2>.

24 For a review, see Tom E Hardwicke et al, 'Calibrating the Scientific Ecosystem through Meta-Research' (2020) 7 *Annual Review of Statistics and Its Application* 11 <https://doi.org/10.1146/annurev-statistics-031219-041104>.

is 'flourishing'[25] by providing insights about how to make the research process more reliable and efficient.[26]

The growth of metaresearch resulted from a perceived 'crisis' in many fields in which studies, peer reviewed and published in eminent journals, could not be replicated by independent researchers.[27] In some cases, researchers were able to reproduce the original methods, but found no effect.[28] In other cases, the studies were simply not reported in enough detail for others to attempt to redo them.[29] In the case of research synthesis, some metaresearch found evidence of bias in that systematic reviews selectively reported outcomes within the articles they included.[30] All of these findings raised serious concerns about research waste and the reliability of existing research evidence.[31] Indeed, in a 2016 survey in *Nature*, 52% of researchers responded that science was experiencing a significant 'reproducibility crisis' and 38% reported it was a slight crisis (only 3% said there was no crisis).[32]

There are several reasons that so many published research findings cannot be replicated, with many of these reasons implicating a lack of transparency and openness. For instance, opaque research reporting allows researchers to leverage undisclosed flexibility in their methods to make results seem more probative of a theory (eg, by excluding some data points after observing their effect on the data).[33] Anonymous surveys find researchers widely engage in this practice, among others (these include criminology, a field that regularly intersects with the legal

25 Munafò et al (n 6) 1.

26 Santo Fortunato et al, 'Science of Science' (2018) 359(6379) *Science* eaao0185:1–7 <https://doi.org/10.1126/science.aao0185>.

27 See Open Science Collaboration, 'Estimating the Reproducibility of Psychological Science' (2015) 349(6251) *Science* 943 <https://doi.org/10.1126/science.aac4716>; Timothy M Errington et al, 'Investigating the Replicability of Preclinical Cancer Biology' (2021) 10 *eLife* e71601:1–30 <https://doi.org/10.7554/eLife.71601>.

28 Open Science Collaboration (n 27).

29 Joscelyn Kaiser, 'Rigorous Replication Effort Succeeds for Just Two of Five Cancer Papers', *Science* (online, 18 January 2017) <https://www.sciencemag.org/news/2017/01/rigorous-replication-effort-succeeds-just-two-five-cancer-papers>. Note that as studies age, the original researchers become unavailable to contact by email to obtain materials: Timothy H Vines et al, 'The Availability of Research Data Declines Rapidly with Article Age' (2014) 24(1) *Current Biology* 94 <https://doi.org/10.1016/j.cub.2013.11.014>.

30 See above n 8. Concerns about the synthesis of medical research often flow from the potential influence of groups like Big Pharma. In the context of law reform, parties (eg, advocacy groups for a particular interest) may seek to produce syntheses that portrays research in a way that influences the result of a law reform project.

31 Iain Chalmers and Paul Glasziou, 'Avoidable Waste in the Production and Reporting of Research Evidence' (2009) 374(9683) *Lancet* 86 <https://doi.org/10.1016/S0140-6736(09)60329-9>; Munafò et al (n 6) 1.

32 Monya Baker, '1,500 Scientists Lift the Lid on Reproducibility' (2016) 533 *Nature* 452 <https://doi.org/10.1038/533452a>. See also John PA Ioannidis, 'Why Most Published Research Findings Are False' (2005) 2(8) *PloS Medicine* e124:1–6 <https://doi.org/10.1371/journal.pmed.0020124>.

33 Joseph P Simmons, Leif D Nelson and Uri Simonsohn, 'False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant' (2011) 22(11) *Psychological Science* 1359 <https://doi.org/10.1177/0956797611417632>.

system).[34] Use of sample sizes that are too small to detect what the researcher is seeking to measure (and failure to plan ahead to use an appropriate sample size) also produces unreliable results.[35] Widespread deficiencies in making research plans, protocols, data, analytic code, and materials openly accessible both allow the above opaque practices to occur and contributes to waste.[36] Finally, incentive structures promote publishing studies with flashy results but poor methods,[37] and a focus on the eminence of researchers over their methods further exacerbates these problems by making peer review less effective.[38]

The response has been a newfound focus on transparent and open practices in both primary research[39] and in synthesis of existing work, sometimes referred to as 'open synthesis'.[40] With respect to open synthesis, this move towards more structured, transparent, and credible summaries of existing work has been described as a 'synthesis revolution'.[41] Fields like psychology and ecology are borrowing and adapting synthesis tools long used in the medical sciences (eg, from the Cochrane Collaboration).[42] This interest in better synthesis is motivated by two reasons central to this article. The first is efficiency (and, accordingly, avoiding waste).[43] If syntheses are not conducted transparently, then they are difficult or impossible to update as new research is conducted.[44] Much of the cost put into the original synthesis is then wasted because it cannot be reused.

Second is the reliability of synthesis. Researchers have raised alarms about syntheses appearing biased in the context of research typically funded by industry (eg, medicine).[45] Synthesis may also be less reliable when it does not

34 See Leslie K John, George Loewenstein and Drazen Prelec, 'Measuring the Prevalence of Questionable Research Practices with Incentives for Truth Telling' (2012) 23(5) *Psychological Science* 524 <https://doi.org/10.1177/0956797611430953>; Hannah Fraser et al, 'Questionable Research Practices in Ecology and Evolution' (2018) 13(7) *PLoS ONE* e0200303: 1–16 <https://doi.org/10.1371/journal.pone.0200303>; Franca Agnoli et al, 'Questionable Research Practices among Italian Research Psychologists' (2017) 12(3) *PLoS ONE* e0172792:1–16 <https://doi.org/10.1371/journal.pone.0172792>; Jason M Chin et al, 'Questionable Research Practices and Open Science in Quantitative Criminology' (2021) *Journal of Quantitative Criminology* (advance) <https://doi.org/10.1007/s10940-021-09525-6>; Matthew C Makel et al, 'Both Questionable and Open Research Practices Are Prevalent in Education Research' (2021) 50(8) *Educational Researcher* 493 <https://doi.org/10.3102/0013189X211001356>.

35 Ioannidis (n 32). While, as Ioannidis notes, many (possibly most) published studies use inadequate sample sizes to find the effects they are testing, facially 'small' sample sizes are appropriate in many cases: Philip L Smith and Daniel R Little, 'Small Is Beautiful: In Defense of the Small-*N* Design' (2018) 25 *Psychonomic Bulletin and Review* 2083 <https://doi.org/10.3758/s13423-018-1451-8>.

36 Hardwicke et al (n 24).

37 Paul E Smaldino and Richard McElreath, 'The Natural Selection of Bad Science' (2016) 3 *Royal Society Open Science* e160384:1–17 <https://doi.org/10.1098/rsos.160384>.

38 Simine Vazire, 'Our Obsession with Eminence Warps Research' (2017) 547 *Nature* 7.

39 See, eg, Nelson, Simmons and Simonsohn, 'Psychology's Renaissance' (n 20); Munafò et al (n 6).

40 See above n 7.

41 'Synthesis Revolution' (n 5).

42 Ibid 489.

43 Nakagawa et al (n 7) 498–9; Polanin, Hennessy and Tsuji (n 15) 1028.

44 Polanin, Hennessy and Tsuji (n 15) 1028.

45 See above n 8.

incorporate unpublished literature,[46] which may be unpublished because it does not support some extant theory or paradigm (a phenomenon sometimes referred to as publication bias).[47] In turn, improving availability and reliability of research synthesis has a downstream aim of improving policymaking; as one synthesist recently said, 'we cannot completely blame policy makers for not adopting more evidence-based approaches in their decision-making process because they must often rely on dubious, premature, scattered and outdated sources of evidence'.[48]

Although synthesis is most relevant to law reform because it involves reviewing *existing* research and data (a task of much law reform), openness and transparency practices in primary research are also relevant. For instance, and as we will see in Part III, law reform bodies do sometimes commission primary research.[49] Accordingly, we will now discuss five transparent and open practices that improve both synthesis and primary research: (pre)registration, registered reports (prioritising methods over results), open data and analysis code, reporting guidelines, and open peer review (see again, Figure 1 for a quick guide to these topics and others). We focus on these five reforms because they have featured heavily in recent best practices guidelines for researchers and synthesists.[50] They are also practices that other researchers have suggested are important for research that informs policy.[51]

A (Pre)registration

Over the past decade, researchers have increasingly started (pre)registering their work.[52] Preregistration, as it is called in the social sciences (medicine prefers

---

46 Alfredo Sánchez-Tójar et al, 'Meta-Analysis Challenges a Textbook Example of Status Signalling and Demonstrates Publication Bias' (2018) 7 *eLife* e37385:1–26 <https://doi.org/10.7554/eLife.37385.029>.

47 See Annie Franco, Neil Malhotra and Gabor Simonovits, 'Publication Bias in the Social Sciences: Unlocking the File Drawer' (2014) 345(6203) *Science* 1502 <https://doi.org/10.1126/science.1255484>.

48 Nakagawa et al (n 7) 498.

49 *RCIRCSA Final Report* (n 2) 44.

50 BA Nosek et al, 'Promoting an Open Research Culture' (2015) 348(6242) *Science* 1422 <https://doi.org/10.1126/science.aab2374> ('TOP Guidelines'); Matthew J Page et al, 'PRISMA 2020 Explanation and Elaboration: Updated Guidance and Exemplars for Reporting Systematic Reviews' (2021) 372 *British Medical Journal* n160:1–36 <https://doi.org/10.1136/bmj.n160> ('PRISMA 2020'). Some are showing positive early returns, with registered reports being linked to improved research quality: Courtney K Soderberg et al, 'Initial Evidence of Research Quality of Registered Reports Compared with the Standard Publishing Model' (2021) 5 *Nature Human Behaviour* 990 <https://doi.org/10.1038/s41562-021-01142-4>; and preregistration being linked to better reporting: Anne M Scheel, Mitchell RMJ Schijen and Daniël Lakens, 'An Excess of Positive Results: Comparing the Standard Psychology Literature with Registered Reports' (2021) 4(2) *Advances in Methods and Practices in Psychological Science* <https://doi.org/10.1177/25152459211007467>.

51 Pamela R Buckley et al, 'The Role of Clearinghouses in Promoting Transparent Research: A Methodological Study of Transparency Practices for Preventive Interventions' (2022) 23(5) *Prevention Science* 787 <https://doi.org/10.1007/s11121-021-01252-5>; Fernando Hoces de la Guardia, Sean Grant and Edward Miguel, 'A Framework for Open Policy Analysis' (2021) 48(2) *Science and Public Policy* 154 <https://doi.org/10.1093/scipol/scaa067>; Evan Mayo-Wilson, 'Evaluating Implementation of the Transparency and Openness Promotion (TOP) Guidelines: The TRUST Process for Rating Journal Policies, Procedures, and Practices' (2021) 6 *Research Integrity and Peer Review* 9:1–11 <https://doi.org/10.1186/s41073-021-00112-8>.

52 See Kupferschmidt (n 20); Christensen et al (n 20).

'registration'),[53] is a public (or embargoed) outline of an intended topic and protocol.[54] It helps prevent duplication of efforts and the negative effects mentioned above of researchers relying on undisclosed analysis flexibility to make their findings seem more convincing (ie, reporting bias). The importance of preregistration is highlighted by the fact that preregistration of clinical medical research has been legally required for over a decade.[55] And, just as it is important that drug trials provide reliable evidence, the evidence underlying law reform should be just as robust. Along these lines, Beerdsen recently recommended that experiments conducted for litigation should be preregistered and Chin and colleagues suggested the same with the reports from forensic sciences in the criminal context.[56] Note, however, that vague and poorly thought-out preregistrations will likely be less effective because it will be easy to fit results-contingent decisions into these plans.[57]

Preregistration also advances reliable synthesis: 'as a minimum necessary requirement, review authors should create a review protocol that documents the various planned method decisions … and publish it in a public, versioned, time-stamped, online registry'.[58] These methods decisions include how studies will be deemed eligible for the synthesis, how the search will be conducted (eg, on what databases and with what keywords), and, if applicable, any ways in which the studies will be assessed (eg, if a certain type of design will be preferred over another).[59] This transparency is also important for the syntheses that law reform bodies rely on to make their policy recommendations.

Preregistration promotes accountability and corresponding reliability in synthesis by allowing users to determine if changes were made to the protocol in light of what was found. These changes may reveal some bias, for instance if the changes all seem made to include research that supports a particular position. On the other hand, the changes may be easily justified by some complication in the synthesis process that was not anticipated beforehand (eg, there is some study outcome that is relevant to the research question that the synthesists had not been

53 Danielle B Rice and David Moher, 'Curtailing the Use of *Preregistration*: A Misused Term' (2019) 14(6) *Perspectives on Psychological Science* 1105 <https://doi.org/10.1177/1745691619858427>.

54 Brian A Nosek et al, 'The Preregistration Revolution' (2018) 115(11) *PNAS* 2600 <https://doi.org/10.1073/pnas.1708274114>.

55 Dickersin and Chalmers (n 20) 535; Dickersin and Rennie (n 20); Ludovic Trinquart, Adam G Dunn and Florence T Bourgeois, 'Registration of Published Randomized Trials: A Systematic Review and Meta-Analysis' (2018) 16(1) *BMC Medicine* 1, 2 <https://doi.org/10.1186/s12916-018-1168-6>.

56 Edith Beerdsen, 'Litigation Science after the Knowledge Crisis' (2021) 106 *Cornell Law Review* 529 <https://doi.org/10.2139/ssrn.3674258>; Jason M Chin, Bethany Growns and David T Mellor, 'Improving Expert Evidence: The Role of Open Science and Transparency' (2019) 50(2) *Ottawa Law Review* 365, 398–9 <https://doi.org/10.2139/ssrn.3345225>.

57 Stuart Buck, 'Beware Performative Reproducibility' (2021) 595 *Nature* 151.

58 Polanin, Hennessy and Tsuji (n 15) 1029. Accordingly, a lack of preregistration is a red flag. 'When there is no preregistration, readers should consider whether any conclusion reached in a review are therefore questionable; of course, authors should draw attention to this limitation': Blair T Johnson, 'Toward a More Transparent, Rigorous, and Generative Psychology' (2021) 147(1) *Psychological Bulletin* 1, 4 <https://doi.org/10.1037/bul0000317>.

59 Polanin, Hennessy and Tsuji (n 15) 1028–9.

aware of). In any case, these data-contingent changes should be scrutinised, and this scrutiny is almost impossible without preregistration.

As further support for the importance of preregistration in the synthesis context, leading systematic review bodies like the Cochrane and the Campbell Collaboration both require preregistration.[60] They follow a three-step process, with preregistration being the first step, followed by creation of detailed study protocol, which is peer reviewed and often published as a standalone document. In this case preregistration is a separate document from the protocol (a detailed description of intended methods), but in other cases preregistration can involve preparing a full detailed protocol, rather than just an outline of the planned study. In situations like these, when the preregistered protocol is peer reviewed, the process is very similar to what is known as a registered report.

B Registered Reports: Peer Review before Results Are Known

Many journals now review and publish articles as 'registered reports' (over 300 accept them, as of this writing).[61] Registered reports encourage researchers to first submit detailed protocols for peer review before any experiments are conducted; after successful peer review of the protocol, the journals guarantee acceptance of the final report, under the condition that the registered protocol is followed. Registered reports address a cause of unreliable science known as the 'results paradox'.[62] In other words, publication decisions should not be conditioned on results because, among other reasons, this incentivises researchers to force the data to fit a publishable narrative, rather than accurately portray what they found. Under the registered report model, it is the methods, hypothesis and background that are peer reviewed – prior to data collection, and then again after analyses are completed.

Besides shifting the incentive to methods over results, registered reports also allow peer reviewers to help improve the methods before data is collected and it is too late.[63] A recent study found that registered reports outpace standard reports on several measures of methodological quality.[64] This focus on methodological quality also helps ensure that null results are more interpretable and meaningful because a failure to support the hypothesis is less likely to be due to a methodological error. All of these benefits apply equally to registered reports for syntheses, a

60 For the Cochrane Collaboration's process, see 'Registering a New Title', *Cochrane Work* (Web Page) <https://work.cochrane.org/registering-new-title>. Campbell follows a similar process (with a slightly more relaxed initial registration form): 'How to Submit a Proposal', *Campbell Collaboration* (Web Page) <https://www.campbellcollaboration.org/idcg-proposal.html>.

61 For a list of journals that offer registered reports and other details about their processes, see Katie Drax, 'Registered Reports: Peer Review before Results Are Known to Align Scientific Values and Practices', *Centre for Open Science* (Web Page) <https://www.cos.io/initiatives/registered-reports>.

62 Chris Chambers, 'What's Next for Registered Reports?' (2019) 573 *Nature* 187, 187 <https://doi.org/10.1038/d41586-019-02674-6>.

63 In our experience, some minor methodological changes may be allowed if they are strongly justified and approved by the reviewers, and would not introduce bias.

64 Soderberg et al (n 50). They also contain a more realistic percentage of null findings: Scheel, Schijen and Lakens (n 50).

point emphasised by the fact that Cochrane and Campbell Collaboration syntheses follow a registered report process.[65]

C Open Data and Analysis (Scripts) Code, and Data Citation Principles

Researchers are also increasingly making their data accessible on public repositories (subject to ethical requirements) and providing the analysis code (scripts) that allows others to reproduce their results.[66] Open data and code make it easier for other researchers to verify results and build upon them (eg, analysing the data in another way, combining them with other datasets). Such aims are furthered by data citation principles that encourage providing persistent citations directly to the datasets.[67] Author guidelines at several journals (ie, the Transparency and Openness Guidelines ('TOP Guidelines')) require statements about whether data, code, and materials are available.[68]

The National Academy of Sciences and other guiding bodies recommend that data be in a machine-readable format and accompanied by metadata explaining the meaning of the variables.[69] Open data and code have other benefits, such as contributing to a 'democratization of knowledge and a better informed citizenry'.[70] And notably, researchers in developing countries often rely on resources from open science.[71] Open data and code practices also facilitate open synthesis by allowing syntheses of underlying raw data, rather than summary statistics.[72]

D Reporting Guidelines

A widely used and empirically tested way of encouraging preregistration and open data and code is through reporting guidelines.[73] These guidelines are often

---

65 See above n 60.

66 See Christensen et al (n 20).

67 'Data citations should facilitate access to the data themselves and to such associated metadata, documentation, code, and other materials, as are necessary for both humans and machines to make informed use of the referenced data ... Unique identifiers, and metadata describing the data, and its disposition, should persist': M Martone, 'Joint Declaration of Data Citation Principles', *FORCE11* (Web Page, 2014) <https://www.force11.org/datacitationprinciples>.

68 Nosek et al, 'TOP Guidelines' (n 50).

69 Committee on Toward an Open Science Enterprise et al (n 1) 53–4, 121, 209–11; Mark D Wilkinson et al, 'The FAIR Guiding Principles for Scientific Data Management and Stewardship' (2016) 3(1) *Scientific Data* 160018:1–9. Researchers have been slower to adopt these more intricate standards for data sharing: Buck (n 57).

70 Committee on Toward an Open Science Enterprise et al (n 1) 36. As we will see in Part III, this is also a stated goal of law reform.

71 Alma Swan, 'Policy Guidelines for the Development and Promotion of Open Access' (Policy Guidelines, United Nations Educational, Scientific and Cultural Organization, 2012).

72 Antica Culina et al, 'How to Do Meta-Analysis of Open Datasets' (2018) 2(7) *Nature Ecology and Evolution* 1053 <https://doi.org/10.1038/s41559-018-0579-2>.

73 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(1) *Systematic Reviews* 60 <https://doi.org/10.1186/2046-4053-1-60>; SeungHye Han et al, 'A Checklist Is Associated with Increased Quality of Reporting Preclinical Biomedical Research: A Systematic Review' (2017) 12(9) *PloS One* e0183591:1–14 <https://doi.org/10.1371/journal.pone.0183591>.

developed through consensus-based procedure involving many stakeholders.[74] They include items that the community thinks are important to disclose when reporting primary research or a synthesis and are often condensed into a short form (or checklist) to be filled in and submitted alongside a study report. These forms help ensure that, even if the author is aware of the importance of reporting some aspect of the research, that they do not accidentally omit something when it is time to submit their work for review and publication.

When it comes to primary research, reporting guidelines promote open data and code, as well as preregistration, by requiring authors to declare if they have used these practices (as in the aforementioned TOP Guidelines). Some reporting guidelines are specific to certain methodologies. For instance, a reporting checklist for behavioural studies provides reporting requirements for those studies, such as the analysis used to determine the appropriate sample size.[75] Similar checklists exist for survey studies.[76]

In the context of synthesis, guidelines require synthesists to declare whether they have reported key aspects, such as their search strategy, the database(s) they searched, the date the last search was conducted, eligibility criteria for inclusion in the synthesis, and whether the synthesis was preregistered.[77] They also ask researchers to declare whether they have attempted to take into account publication bias, which is a serious challenge for research synthesis. In short, it is easy for syntheses to produce unreliable results because it is easier to find published articles, which are those that tend to support the researchers' hypothesis. This can overstate the support for the finding in the synthesis. Some popular reporting guidelines include: the Preferred Reporting Items for Systematic Reviews and Meta-Analyses ('PRISMA'),[78] Meta-Analysis Reporting Standards ('MARS'),[79] RepOrting Standards for Systematic Evidence Syntheses ('ROSES'),[80] and the reporting guidelines of the Cochrane and Campbell Collaborations.[81] Similarly,

74 Aczel et al (n 22) used such a procedure. Although note that some guidelines have been criticised for not being developed using a consensus-based procedure: Masahiro Banno, Yasushi Tsujimoto and Yuki Kataoka, 'The Majority of Reporting Guidelines Are Not Developed with the Delphi Method: A Systematic Review of Reporting Guidelines' (2020) 124 *Journal of Clinical Epidemiology* 50 <https://doi.org/10.1016/j.jclinepi.2020.04.010>.

75 Aczel et al (n 22).

76 'Survey Disclosure Checklist', *American Association of Public Opinion Research* (Web Page, April 2021) <https://www.aapor.org/Standards-Ethics/AAPOR-Code-of-Ethics/Survey-Disclosure-Checklist.aspx>. See also Ines Schaurer, Tanja Kunz and Tobias Heycke, 'Documentation of Online Surveys' (Guideline, April 2020).

77 David Moher et al, 'Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement' (2009) 6(7) *PloS Medicine* e1000097:1–6 <https://doi.org/10.1371/journal.pmed.1000097>; Neil R Haddaway et al, 'ROSES RepOrting Standards for Systematic Evidence Syntheses: Pro Forma, Flow-Diagram and Descriptive Summary of the Plan and Conduct of Environmental Systematic Reviews and Systematic Maps' (2018) *Environmental Evidence* 7:1–8 <https://doi.org/10.1186/s13750-018-0121-7> ('ROSES').

78 Moher et al (n 77). PRISMA was updated in March 2021: Page et al, 'PRISMA 2020' (n 50).

79 'American Psychological Association (APA) Meta-Analysis Reporting Standards (MARS)' (Reporting Standards, American Psychological Association, 2015).

80 Haddaway et al, 'ROSES' (n 77).

81 See above n 60.

other organisations have developed standards for assessing existing syntheses to determine how thoroughly they were conducted.[82]

Reporting standards like PRISMA are important to the transparency of synthesis: 'reviews of scientific evidence cannot be recognized as having high quality without reporting all the relevant details involved in completing the review'.[83] Non-transparent syntheses are wasteful because they are difficult or impossible to update, equating to 'research and resource waste to the highest degree, akin to collecting data from thousands of individuals, reporting some summary statistics, and then throwing away the data set'.[84] Future synthesists will not know how studies were or were not included in the previous synthesis and so they will have to both redo that work on top of their own update. On the other hand, syntheses that accord with PRISMA and its ilk enable cumulative work; researchers can follow the previous protocol to add newer research. This allows us to see a fuller picture of the existing knowledge and explore new trends that may emerge.

E Open Peer Review

Finally, many journals, which publish primary research and syntheses, operate on an open peer review model (sometimes called 'open reports') in which the peer reviews are published along with the articles.[85] This increases the transparency of the peer review and publishing process by allowing the readers to see for themselves what the reviewers thought were the strengths and weaknesses of the article, and how the authors responded.[86]

III RECENT CONCERNS ABOUT EVIDENCE-BASED REFORM

Modern law reform bodies often state preferences for evidence-based reform.[87] These can be found in mission statements of standing bodies,[88] reflections on the

82 Beverley J Shea et al, 'AMSTAR 2: A Critical Appraisal Tool for Systematic Reviews That Include Randomised or Non-Randomised Studies of Healthcare Interventions, or Both' (2017) 358 *British Medical Journal* j4008:1–9 <https://doi.org/10.1136/bmj.j4008>; Paul Woodcock, Andrew S Pullin and Michael J Kaiser, 'Evaluating and Improving the Reliability of Evidence Syntheses in Conservation and Environmental Science: A Methodology' (2014) 176 *Biological Conservation* 54 <https://doi.org/10.1016/j.biocon.2014.04.020>.

83 Johnson (n 58) 1.

84 Polanin, Hennessy and Tsuji (n 15) 1028.

85 Tony Ross-Hellhauer, 'What Is Open Peer Review? A Systematic Review' (2017) 6 *F1000Research* 588:1–37 <https://doi.org/10.12688/f1000research.11369.1>.

86 Birgit Schmidt et al, 'Ten Considerations for Open Peer Review' (2018) 7 *F1000Research* 969:1–12 <https://doi.org/10.12688/f1000research.15334.1>.

87 See above n 2.

88 For example, the ALRC states that it 'supports the Attorney-General and the Australian Government by providing *evidence-based research* to inform government decisions about the development, reform and harmonisation of Australian laws and related processes': *ALRC Annual Report* (n 2) 2 (emphasis added). The Law Commission for England and Wales states that '[t]he work of the Commission is *grounded in thorough research* and analysis of case law, legislation, academic and other writing, and other relevant sources of information both in the UK and overseas': *LCEW Annual Report* (n 2) 57 (emphasis added).

law reform process from law reform commissioners,[89] and in reports about specific projects.[90] For instance, the Law Commission for England and Wales provides a memo laying out a detailed statement about its use of research: 'From its earliest days, therefore, the Commission has had regard to the findings of both qualitative and quantitative research. … These exercises have provided an essential supplement to our findings from our own consultation.'[91]

Statements like these are consistent with what appears to be an increasing interest among governments in evidence-based policy.[92] And researchers, for their part, are often eager to provide work that will have policy consequences.[93] For Australian researchers, such engagement is encouraged by universities and funding bodies.[94]

Despite stated commitments to evidence-based policy, fulfilling this goal often attracts criticism in practice. We will now discuss some of these critiques of both the research undertaken by law reform bodies themselves and that which they hire from outside researchers. We end this part with a preliminary study of the research synthesis reports commissioned by the *RCIRCSA*.

A Evidence Synthesis Undertaken by Law Reform Bodies

While the importance of evidence-based policy is not contentious, how it works in practice sometimes is. Law reform bodies have drawn criticism for the way in which they collect research, weigh up the evidence that is presented to them and eventually prefer some pieces of evidence over others.[95] As we will see, open synthesis and science cannot resolve value judgments about whether, for instance, quantitative studies should be preferred over the individuals' lived experiences.

89 'The Commission sees cooperation with other agencies engaged in critical reflection about Canadian law as a way to avoid duplication of effort. Forging networks and building research partnerships with independent scholars and public policy organizations is one strategy to generate momentum for legal change': Macdonald (n 4) 105 (emphasis omitted), but note that Canada's commission later closed. See also Partington (n 4) 382–3.

90 'The value of reliable, complete, and accessible data in developing rational and effective law and policy cannot be understated': Australian Law Reform Commission, *Corporate Criminal Responsibility* (Final Report No 136, April 2020) 118 ('*Corporate Criminal Responsibility Report*'). See also *RCIRCSA Final Report* (n 2) 42.

91 Law Commission for England and Wales, 'The Law Commission for England and Wales and Its Use of Empirical Research' (Research Note, 2009) <https://s3-eu-west-2.amazonaws.com/lawcom-prod-storage-11jsxou24uy7q/uploads/2015/06/empirical_research_090610.pdf>.

92 Miguel et al (n 2) 30; 'Collaborative Assessment for Trustworthy Science: The RepliCATS Project', *The University of Melbourne* (Web Page) <https://replicats.research.unimelb.edu.au/>.

93 Christopher Uggen and Michelle Inderbitzin, 'Public Criminologies' (2010) 9(4) *Criminology and Public Policy* 725 <https://doi.org/10.1111/j.1745-9133.2010.00666.x>; Nakagawa et al (n 7) 498; Polanin, Hennessy and Tsuji (n 15) 1028.

94 See Katy Barnett, 'Citation as a Measure of "Impact": Female Legal Academics at a Disadvantage?' (2019) 44(4) *Alternative Law Journal* 267 <https://doi.org/10.1177/1037969X19874847>.

95 Kay Cook and Kristin Natalier, 'Gender and Evidence in Family Law Reform: A Case Study of Quantification and Anecdote in Framing and Legitimising the "Problems" with Child Support in Australia' (2016) 24 *Feminist Legal Studies* 147 <https://doi.org/10.1007/s10691-016-9317-9>; Bri Lee, 'The Old Guard Preventing Reform to Consent Laws', *The Saturday Paper* (online, 15 August 2020); 'BREAKING: Queensland Law Reform Commission Fails to Make Substantive Recommendations to Improve Rape Law', *Rape and Sexual Assault Research and Advocacy* (Web Page) <https://rasara.org/qlrc>.

But, these practices can still provide considerable and previously untapped utility by clarifying the process by which evidence is gathered and balanced, and by helping control biases that creep into the human processes at play in law reform.

Open synthesis may have assisted with the 2020 Queensland Law Reform Commission ('QLRC') report that refrained from recommending substantial strengthening to sexual assault laws. Commentators and advocacy groups strongly criticised the report, saying the QLRC 'cherry-picked' one United Kingdom ('UK') study that found rape myths were not widespread.[96] These parties also noted that the QLRC did not seem to thoroughly survey existing research.[97] From a review of the report, it indeed appears that the QLRC's review of research on rape myths (the degree to which they are held and whether they affect verdicts) is unstructured and unclear as to its scope.[98] It cites a handful of studies without explaining how they were identified. Ultimately, it seems to prefer the UK study because it polled actual jurors, rather than a sample of people voluntarily responding to a survey.[99]

Similarly, Cook and Natalier reviewed transcripts from the controversial 2003 Parliamentary Inquiry into Child Custody Arrangements, finding that 'both anecdote and empirical data are criticised and praised at different times'.[100] This uneven weighting of evidence raised the possibility of bias because quantitative evidence was disregarded when it came to findings that challenged the inquiry's ultimate conclusion. In particular, the chair of the inquiry at one point said: 'I am a bit of an anti-research person … it is time we get out of the research and get into delivering exactly what our families need.'[101] However, later in the proceedings, she seemed to require quantitative research: 'it is imperative that independent modelling of the cost of children in separated families should be undertaken and published.'[102] These statements raised the possibility that narrative and quantitative evidence were being selectively preferred to reach the inquiry's preferred outcome, despite both being valuable evidence.

As we will develop in Part IV, the experience with the QLRC report and the inquiry into child custody arrangements demonstrate the need for open synthesis in law reform. Commentators were seeking clarity about why some research was included in the report and why some was not. Under the open synthesis model, synthesists provide a table of sources with reasons for exclusion. These reasons can then be compared to the preregistration that describes the scope (ie, inclusion and exclusion criteria) that was determined before looking at the research. This would help illuminate and control biases that may have produced a change of approach that occurred once the available research was assessed. Moreover, when it comes to weighing up research, there may be reasons to value one type of study over another (eg, foreign versus domestic, actual jurors versus mock jurors), or narratives over

96 Lee (n 95); Rape and Sexual Assault Research and Advocacy (n 95).
97 Lee (n 95).
98 Queensland Law Reform Commission, *Review of Consent Laws and the Excuse of Mistake of Fact* (Report No 78, June 2020) xi–xii, 206–8.
99 Ibid 208, 210.
100 Cook and Natalier (n 95) 148.
101 Ibid 149.
102 Ibid.

quantitative research. But, when that preference is stated *after* seeing that data, the potential for conscious or unconscious bias is high. A preregistration helps ensure that the authors will disclose and confront their selective use of evidence.

Finally, a firsthand perspective comes from the authors and researchers behind the 2020 ALRC report on corporate criminal responsibility.[103] The report explicitly endorsed the importance of evidence-based reform: 'The value of reliable, complete, and accessible data in developing rational and effective law and policy cannot be understated'.[104] At the same time, the authors observed they struggled to find the data they were seeking, noting considerable '[d]ifficulties in accessing quantitative data relating to corporate criminal responsibility'.[105] The authors went on to collect their own data as provided by the Commonwealth Director of Public Prosecutions, the Australian Securities and Investments Commission, the New South Wales Bureau of Crime Statistics and Research, the Australian Bureau of Statistics, and others.[106] The report authors also provided a large data appendix that further detailed their methods and contained several tables of relevant statutes, regulations, and legal decisions their staff had coded.[107]

Despite the ALRC report emphasising the importance of accessible data and the admirable amount of work that went into the process, their report falls short of many of the practices we reviewed in Part II.[108] Data files are not provided at any point. The raw data that the report does provide are in PDF tables that would be very difficult to extract, reuse, and combine with other data. When data from external sources is cited, it is not to the specific document. Rather, for example, the ALRC cites a customised report from the Australian Bureau of Statistics ('ABS') with simply a link to the general website of the ABS[109] or as general correspondence to the ALRC.[110] The analytic code that produces the summary statistics, tables, and figures is not provided. This code is important because of the complexity of the ALRC's processes that produced the statistics and tables, any details of which are often embedded in footnotes to cells of the table.[111] In other words, providing the code would allow researchers and citizens to better understand how the statistics were calculated and the assumption that go into that process. Finally, the entirety of the data collection and analysis endeavour was not preregistered, opening the ALRC up to allegations that it selectively included and excluded data, analytic approaches, and results.

---

103 *Corporate Criminal Responsibility Report* (n 90) 70, 118, 204.
104 Ibid 118.
105 Ibid 117.
106 Ibid 70–3.
107 Australian Law Reform Commission, *Corporate Criminal Responsibility: Data Appendices* (Report No 136, April 2020).
108 Committee on Toward an Open Science Enterprise et al (n 1) 121; Wilkinson et al (n 69); Martone (n 67).
109 *Corporate Criminal Responsibility Report* (n 90) 72 n 10. See also *Corporate Criminal Responsibility Report* (n 90) 72 n 13.
110 Ibid 71–2 nn 3, 4, 11, 14. Compare these citations to the Force11 Data citation principles, Martone (n 67).
111 For instance, see the footnote to cell 6 of Table 3-1. 'This figure is not explicit in the relevant underlying data. It has been derived by reference to the stated number of successful prosecutions': *Corporate Criminal Responsibility Report* (n 90) 101 n 152.

To be clear, we are not suggesting any negligence or malfeasance. The ALRC is staffed by lawyers and those who are in law school, or who just finished law school; they cannot be expected to have training in methodology in research synthesis (see Part IV).[112] Moreover, the general approach to research may be different for lawyers, who are often trained to research legal authorities that will support a particular position. Still, given the huge amount of work that went into the report and research, it is unfortunate that open practices were not followed to facilitate reusability, reanalysis, and reliability.

B Research Funded and Commissioned by Law Reform Bodies

Given the inherent challenges in asking law reform staff to conduct their own research, be that synthesis or gathering new data, an attractive alternative is to commission research from universities and other research institutes. This option was heavily relied on by the well-funded *RCIRCSA* (although they had in-house research staff as well).[113] They commissioned many primary research studies (eg, surveys, focus groups) and syntheses. We will focus on the *RCIRCSA*'s widely-discussed mock jury study and then turn to the syntheses it commissioned.[114]

The *RCIRCSA* mock jury study aimed to address a gap in the literature (although note that the gap was determined based on a synthesis that lacked transparency: see below footnotes 128–32 and accompanying text),[115] about whether there was a 'joinder effect'.[116] In other words, they tested whether mock juries in child sexual assault cases decide differently in joint trials (ie, accusations by multiple complainants are heard all at once) versus trials with one complainant (but supported by evidence from other witnesses as to the accused's tendency towards child sexual assault).[117] The presence of a joinder effect would raise the risk of unfair prejudice by hearing child sexual assault cases as joint trials. Mock jurors read scenarios describing these different types of trials. Juries convicted more in the joint trial condition, but the difference was not statistically significant (ie, it could not be distinguished from random variation based on a commonly used method and cut-off; more on this below).[118] The authors concluded that 'no joinder effect was found'.[119] This conclusion appears to have been important in the *RCIRCSA*'s decision to recommend loosening the rules that result in

112 In fact, compliance with preregistration, open data, and many other reforms is uneven among researchers outside law reform: Buck (n 57).

113 'RCIRCSA Research Website' (n 14).

114 Hunter and Kemp (n 14); Jane Goodman-Delahunty et al, 'Response to Submissions to the Royal Commission's Consultation Paper in Relation to Jury Reasoning Research' (Response Paper, December 2016) <https://www.childabuseroyalcommission.gov.au/sites/default/files/research_report_-_jury_reasoning_research_-_request_for_feedback_and_response_from_researchers_-_government_responses.pdf> ('Mock Jury Study Responses').

115 Jane Goodman-Delahunty, Annie Cossins and Natalie Martschuk, 'Jury Reasoning in Joint and Separate Trials of Institutional Child Sexual Abuse: An Empirical Study' (Research Report, Royal Commission into Institutional Responses to Child Sexual Abuse, May 2016) 71.

116 Ibid.

117 Ibid 184. For brevity, we will gloss over much of the nuances of the study that are unrelated to transparency and research design, but for a brief summary: see Hunter and Kemp (n 14) 257–9.

118 Goodman-Delahunty, Cossins and Martschuk (n 115) 184.

119 Ibid 94.

joint trials.[120] Their recommendation was taken up, with New South Wales loosening the rules that allow joint trials in child sexual assault case.[121] The mock jury study has also been cited by several courts.[122]

In a subsequent article, Hunter and Kemp discussed an important limitation of the *RCIRCSA* study, that the lack of a statistically significant joinder effect belied the fact that if one additional jury in the joint trial condition had convicted, the effect would have been statistically significant.[123] This observation is related to the broader issue we discussed in Part II, which is that sample sizes must be properly considered prior to the study being conducted.[124] In the context of the *RCIRCSA*, this means that if the experimenters did not study enough juries to find what they were looking for, then the absence of a finding is meaningless (ie, they did not have enough 'statistical power' to find a joinder effect that matters in the world). Here, the widely used analogy to a telescope may be of use.[125] If one looks at a particular part of the sky with a telescope that is not strong enough to detect an astral body of a given size and distance, then the observer cannot validly claim that there is no such body there. It might just be that the telescope is not strong enough.

Given the importance of statistical power, methodologists have long exhorted researchers in applied settings to determine and preregister *prior* to running the study what the smallest effect size of interest is, and how many participants will be needed to find that.[126] Those calculations should be transparently disclosed. For instance, in the current context, what is the smallest meaningful difference between joint and non-joint trials? If the experimenters and law reformers think that, say, no greater than 10% more convictions in joint trials is acceptable, then they should include enough mock jurors to have the statistical power to find that 10% effect. This would give meaning to a lack of a statistically significant effect. Going back to the telescope analogy, researchers would have to state, prior to looking at the sky with a telescope, the smallest planet of interest at a certain distance and then pick a telescope powerful enough to find that planet. Only then, if they see no planet, it

120 Hunter and Kemp (n 14) 253.

121 *Evidence Act 1995* (NSW) s 97A.

122 See, eg, *Vojneski v The Queen* (2016) 262 A Crim R 370; *R v Kisun* [2016] ACTSC 212.

123 '[T]his masks the fact that the data provide stronger support for the hypothesis that there is a joinder effect than for the null hypothesis that there is no joinder effect (a likelihood ratio of 3.58). In fact, a change of decision by just one jury (eg 11/17 in the tendency evidence trials or 13/16 in the joint trials) would have been sufficient to render this a statistically significant result': Hunter and Kemp (n 14) 258.

124 Ioannidis (n 32).

125 Uri Simonsohn, 'Small Telescopes: Detectability and the Evaluation of Replication Results' (2015) 26(5) *Psychological Science* 559 <https://doi.org/10.1177/0956797614567341>.

126 Anne G Copay et al, 'Understanding the Minimum Clinically Important Difference: A Review of Concepts and Methods' (2007) 7(5) *Spine Journal* 541 <https://doi.org/10.1016/j.spinee.2007.01.008>. For a review, see Daniël Lakens, 'Sample Size Justification' (Research Paper, Eindhoven University of Technology, January 2021). Note that the mock jury study authors do not provide their power calculations. Rather, they simply say: 'the effect size was determined based on the magnitude of effects observed in past studies': Goodman-Delahunty, Cossins and Martschuk (n 115) 20. They do not say what this effect size is, nor do they provide their calculations. More importantly, their approach is nonsensical because their stated reason for conducting the study was the lack of an analogous joinder effect study (ie, the gap they found in the literature). How then could they use an existing effect size to determine the size of an effect they were likely to find?

is valid to infer that one of the specified size and distance is not there. As it stands in the *RCIRCSA* mock jury study, the fact that the experimenters did not find an effect is meaningless, and so is any reform that followed from that conclusion.[127]

As noted above, the *RCIRCSA* mock jury study was commissioned to address a gap in the literature. However, it is not clear whether there really was a gap because of the lack of transparency of the syntheses conducted. A rigorous and transparent systematic evidence 'gap map' is the method of choice to inform research priorities.[128] Systematic evidence maps are similar to systematic reviews of evidence in adapting systematic ways to collect and collate evidence, with clear a priori inclusion and exclusion criteria. However, while systematic reviews synthesise the actual evidence within studies, maps focus on the presence or absence of the studies providing the evidence and the broader patterns of how evidence is distributed.[129]

It is worth noting other non-transparent elements of the mock jury study that hinder its long-term utility. First, the data is not available and so it cannot be reanalysed and combined with future research. The analysis scripts are also not available and the study was not preregistered (as it would have been if it was a clinical trial – although the public significance of this study that directly informed a new criminal evidence law is arguably on par with medical research). While this study was peer reviewed by 'three eminent law and social science academics' after it was conducted, as the authors note in their response to submissions to the *RCIRCSA* critiquing the study,[130] these were not disclosed (ie, it was not open peer review). Moreover, peer review was not conducted prior to the data being collected such that it could actually inform how the study was conducted. This would have been the case if it was a registered report, which tend to be of higher methodological quality.[131] Finally, the *RCIRCSA* was criticised on the grounds that it selectively cited previous research.[132] There may have been good grounds for what studies the researchers did uncover, but because there was no transparent gap analysis, the criticisms land more forcefully.

Finally, we emphasise that while the results of the mock jury study were ambiguous (recall, one flipped jury decision could have swayed the results), post hoc conjecturing about possible results does not help with the larger question of whether there is a joinder effect. To address that, researchers should conduct a preregistered replication with a large enough sample such that no effect is found,

---

127 Another reason metaresearchers are concerned with small effect sizes is that they lead to poor estimates of effect sizes, raising the possibility of confusing a small effect for a large one: Andrew Gelman and John Carlin, 'Beyond Power Calculations: Assessing Type S (Sign) and Type M (Magnitude) Errors' (2014) 9(6) *Perspectives on Psychological Science* 641 <https://doi.org/10.1177/1745691614551642>.

128 Ashrita Saran and Howard White, 'Evidence and Gap Maps: A Comparison of Different Approaches' (2018) 14(1) *Campbell Systematic Reviews* 1 <https://doi.org/10.4073/cmdp.2018.2>.

129 Isomi M Miake-Lye et al, 'What Is an Evidence Map? A Systematic Review of Published Evidence Maps and Their Definitions, Methods, and Products' (2016) 5 *Systematic Reviews* 28:1–21 <https://doi.org/10.1186/s13643-016-0204-x>.

130 'Mock Jury Study Responses' (n 114) 1.

131 Soderberg et al (n 50).

132 'Mock Jury Study Responses' (n 114) 24.

then it can be safely concluded that a meaningful effect does not exist. This is especially important (as is replication work more generally) for work that drives policy. In such cases, we want to know not just whether there is an effect, but if we do not find that effect (ie, a null result), we need to know how to interpret that non-finding.

C An Exploratory Study of *RCIRCSA* Synthesis Reports

To further explore the openness and transparency of research commissioned by law reform, we leveraged the large number of research reports commissioned by the *RCIRCSA*. These reports provide a distinctive glimpse into the level of synthesis transparency being provided to a well-funded royal commission, whose recommendations have already affected the law. Given the expense of the process and the importance of the subject matter, full transparency should have been strongly encouraged by the *RCIRCSA*. The reports covered important topics, such as studies on the effects of institutional sexual abuse on survivors and studies of the effectiveness of specialist prosecution units and courts for child sexual offences.[133]

We coded the 48 research reports contained on the *RCIRCSA*'s research website.[134] As described in our preregistered protocol,[135] we limited our analysis to syntheses (several reports include primary research, like surveys and interviews), of which there were 22. We coded several characteristics of synthesis that widely-used reporting guidelines and synthesis assessment measures also examine.[136] However, for efficiency and because we focused on transparency, we coded only a truncated set of measures: (1) whether the protocol was preregistered; (2) whether the protocol was peer reviewed prior to data collection; (3) whether a reporting guideline was used; (4) whether all search terms were reported; (5) whether all databases searched were reported; (6) whether the grey literature (eg, work not published in traditional academic journals) was searched; (7) whether all inclusion criteria were clearly described; and (8) whether excluded material was listed along with the reason it was excluded. These are also items that other recent assessments and theoretical accounts of policy-focused research have studied.[137] Our coding scheme is detailed in our preregistration.[138] We also coded whether any one of the authors had a specific specialisation in information research (ie, worked at research centre or library) as indicted in the report and without doing any further research

133 Tamara Blakemore et al, 'Impacts of Institutional Child Sexual Abuse: A Rapid Review of the Evidence' (2017) 74 *Child Abuse and Neglect* 35; Nina Westera, Elli Darwinkel and Martine Powell, 'A Systematic Review of the Efficacy of Specialist Police Investigative Units in Responding to Child Sexual Abuse' (Research Report, Royal Commission into Institutional Responses to Child Sexual Abuse, March 2016).

134 'RCIRCSA Research Website' (n 14).

135 Jason Chin, Malgorzata Lagisz and Shinichi Nakagawa, 'Preregistration', *OSF HOME* (Web Page, 2021) <https://osf.io/2wp75/registrations>.

136 See PRISMA 2020 (n 50); Haddaway et al, 'ROSES' (n 77); American Psychological Association (n 79).

137 See Buckley et al (n 51); de la Guardia, Grant and Miguel (n 51); Mayo-Wilson (n 51). We also drew from the coding scheme of an ongoing assessment of syntheses: Matthew J Page et al, 'The REPRISE Project: Protocol for an Evaluation of REProducibility and Replicability in Syntheses of Evidence' (2021) 10 *Systematic Reviews* 112:1–13 <https://doi.org/10.1186/s13643-021-01670-0>.

138 Ibid. See also Jason Chin, Malgorzata Lagisz and Shinichi Nakagawa, 'Coding Scheme Metadata and Further Details on Methods', *OSF HOME* (Web Page, 2021) <https://osf.io/edaz5/>.

into their backgrounds. Our raw data and code are also publicly available.[139] Given the small sample sizes and our lack of an a priori hypotheses, this study should be seen as exploratory and descriptive.

Table 1: The transparency and openness of *RCIRCSA* synthesis reports

A summary of the transparency and openness of the 22 research syntheses commissioned by the RCIRCSA. The second column summarises all of the reports. The third and fourth row break this down into reports not authored by information research specialists (n = 15) and those that were authored by such individuals (n = 7), respectively. Cells are the percentage of reports in that column that meet the transparency and openness guideline (rounded to whole numbers). See Figure 2 for an infographic displaying this data. The code that produces the data in this table is available online, as is the codebook that further describes how these reports were coded and the meaning of the rows: <https://osf.io/2wp75/>.

	All reports (N = 22)	**Non-specialist (n = 15)**	**Specialist (n = 7)**
Preregistered	9%	7%	14%
Peer-reviewed protocol	0%	0%	0%
Reporting guideline cited	18%	7%	43%
Search terms reported	77%	67%	100%
Databases reported	64%	47%	100%
Grey literature searched	62%	57%	71%
Inclusion criteria clearly stated	59%	40%	100%
Excluded articles provided with reasons	5%	0%	14%

Overall, we found an extremely uneven level of synthesis transparency (Table 1 and Figure 2). Starting with the more promising results, reporting all search terms, databases searched, and inclusion criteria were relatively common (60–75% of syntheses met these standards). Still, it is somewhat troubling that so many syntheses failed to report these fundamental aspects of their methodology. On that note, only about 10% (2/22) of syntheses were preregistered, none of those preregistrations were peer reviewed (or, at least, this was not declared), and 18% (4/22) used a reporting guideline (PRISMA, in all cases). Only one synthesis listed excluded articles with reasons for exclusion. We also note that among the two preregistered syntheses, the authors simply noted that they had pre-specified their protocol but did not provide a link to that preregistration. This falls well short of practices that would allow the reader to compare the preregistered protocol to the one used.

139 Jason Chin, Malgorzata Lagisz and Shinichi Nakagawa, 'Raw Data', *OSF HOME* (Web Page, 2021) <https://osf.io/yc34f/>; Jason Chin, Malgorzata Lagisz and Shinichi Nakagawa, 'Code', *OSFHOME* (Web Page, 2021) <https://osf.io/4bakn/>.

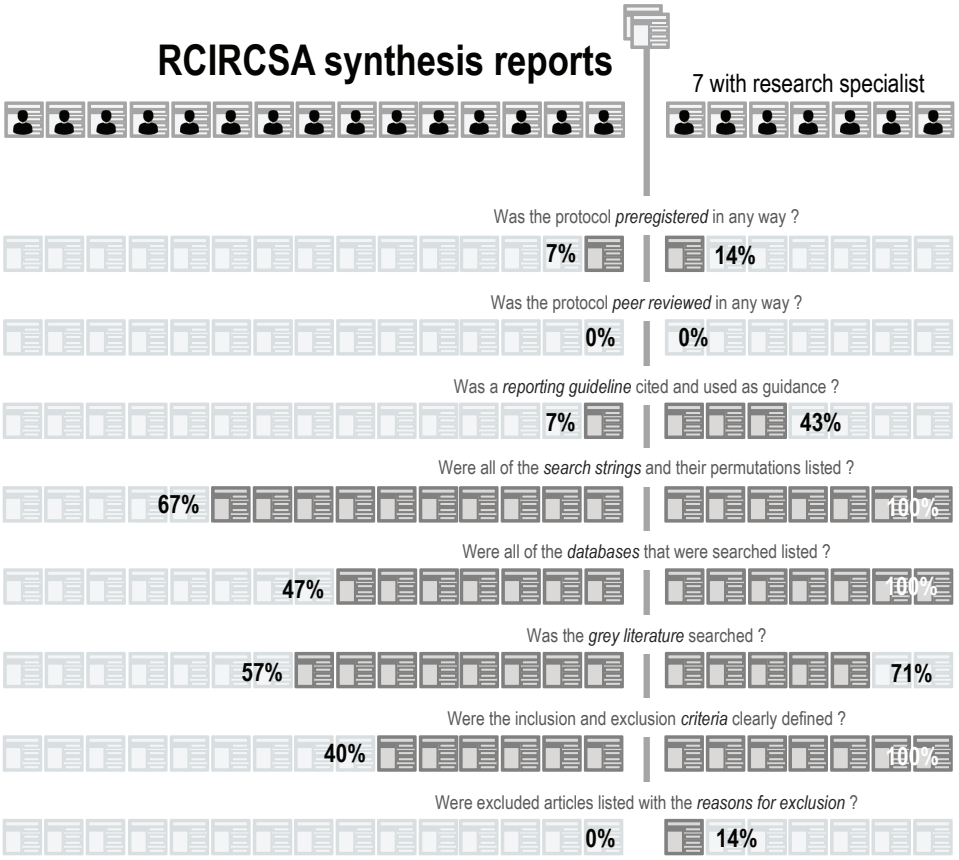

Figure 2: The transparency and openness of *RCIRCSA* synthesis reports, infographic

An infographic displaying the results in Table 1. The 15 reports on the left side of the graphics are reports with teams that did not include an author with some affiliation with a research centre or library. The seven on the right side did include such an author.

Although we again caution against drawing too strong an inference given the small sample, we also point out that syntheses with an information research specialist (7 of the 22 syntheses) performed better than those without one on all metrics (see Table 1 and Figure 2). Almost half of such reports used a reporting guideline, and all of them fully reported their search terms, databases searched, and inclusion criteria.

While we focused on syntheses, it is also worth noting that the 18 reports that described primary research missed opportunities to make their work verifiable and reusable.[140] For instance, reports did not provide the analytic code that produced

140 See these in our datafile, Chin, Lagisz and Nakagawa, 'Raw Data' (n 139).

their statistics.[141] None of them were preregistered. And, while there were likely good reasons not to in most studies (eg, privacy), none provided their data.[142] In any case, the widely used TOP Guidelines, which apply to journal articles, would still require that researchers explain why they have not made their data open access.[143] We cannot think of a good reason why the same standard should not apply to the publicly funded knowledge generated by law reform bodies.

We also found that many reports (not included in Table 1 or the description above) were syntheses of statutes, judicial decisions, and policies. Metaresearch has not focused on this type of legal synthesis and reporting guidelines do not explicitly address them. Still, many of the same concerns with opaque synthesis apply to such work. For instance, researchers could leave out certain laws or policies on an ad hoc basis to provide a biased view of the state of the law. Future research may wish to establish best practice guidelines for this legal synthesis.

IV FIVE SUGGESTIONS FOR OPEN LAW REFORM

> '*The greater the diversity and openness of a community and the stronger its protocols for supporting free and open debate, the greater the degree of objectivity it may be able to achieve as individual biases and background assumptions are "outed," as it were, by the community.*'[144]

Metaresearch findings and concomitant changes to the research process in fields like psychology and ecology ought to prompt similar introspection among law reform bodies. Indeed, these changes (eg, preregistration, open data) are moored in many of the same principles that law reform bodies endorse, such as open access, transparency, and democratisation of information.[145] And, both seek to fulfil the same mission: more reliable outcomes and more efficient processes. Starting with the most immediately achievable outcomes and move into more ambitious and long-term projects, we will map out a path towards open law reform. Our proposed model is summarised in Figure 3. The proposed ('Future') model incorporates the open science and syntheses practices discussed in Part II: (pre)registration of report plans and protocols, registered reports (ie, early-stage peer review of methodology), open

141 For instance, despite no apparent privacy issues in the majority of sentencing decisions analysed (eg, offender's age, level of court, sentencing date, offence) there was no open data and code in Karen Gelb, *A Statistical Analysis of Sentencing for Child Sexual Abuse in Institutional Contexts* (Research Report, Royal Commission into Institutional Responses to Child Sexual Abuse, March 2016) 47. Similarly, open analytic code would have rendered the results verifiable in Tim Moore et al, 'Our Safety Counts: Children and Young People's Perceptions of Safety and Institutional Responses to Their Safety Concerns' (Research Report, Royal Commission into Institutional Responses to Child Sexual Abuse, September 2016).

142 However, note that the data from many variables from sentencing decisions (which are often publicly available) could have been made open access (along with the code that produces summary statistics from that data), see Gelb (n 141).

143 Nosek et al, 'TOP Guidelines' (n 50).

144 Naomi Oreskes, *Why Trust Science?* (Princeton University Press, 2019) 53.

145 Macdonald (n 4) 104–5; Partington (n 4) 391–2; *Royal Commission into Family Violence* (n 4) 2.

data and scripts, encouragement of these practices through reporting guidelines, and open peer review. These goals can be advanced through methodologically diverse teams (see below footnotes 154–9 and accompanying text).

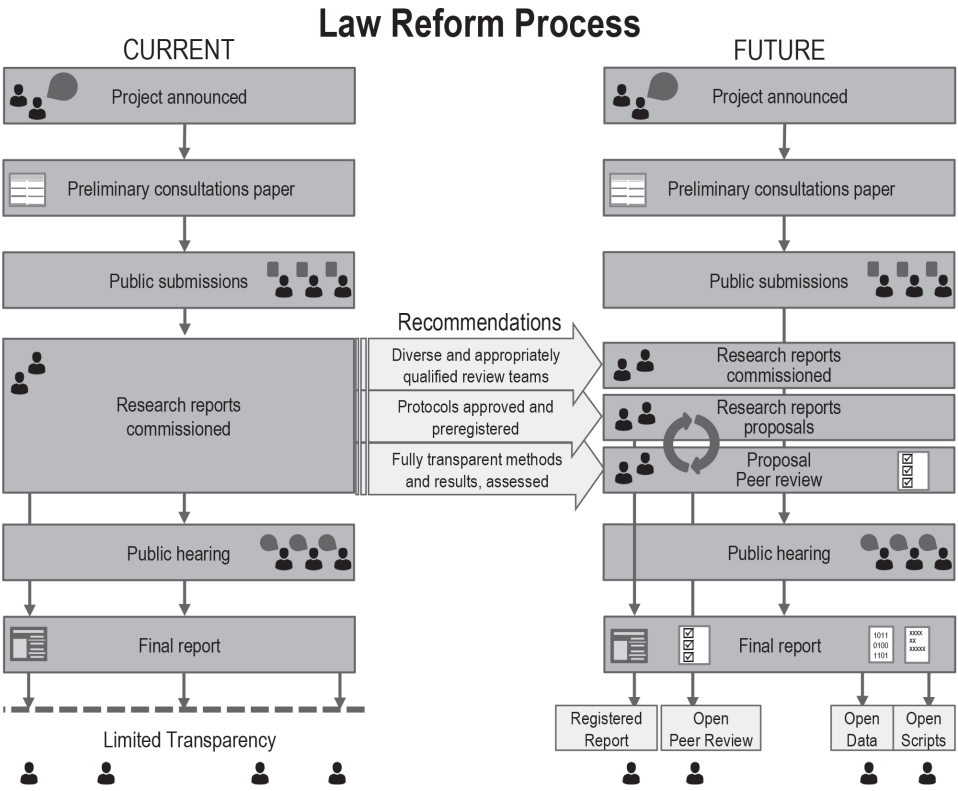

Figure 3: A plan for open law reform

The standard law reform model (current) compared to an open and transparent evidence-based law reform process (future).

First, law reform staff should require that researchers conducting syntheses for them follow standard reporting guidelines (eg, PRISMA, ROSES, MARS). This will help achieve a standard of quality and transparency that is set by the law reformers and not the contracted researchers. As we have seen, it is not realistic to assume that researchers will, of their own volition, conduct open synthesis. It will also help mitigate any biases that the researchers bring into the process (eg, their preferred policy outcome).

In situations where time and cost are an especially limiting factor, the law reform body and researcher can use a synthesis format that takes these limits into account, but that still requires full transparency. These are sometimes referred

to as rapid reviews, or in other words a 'systematic review with shortcuts'.[146] We saw four synthesis reports in our *RCIRCSA* review that used some type of a rapid review format.[147] Unfortunately, these reviews did not take simple steps towards transparency, such as preregistration, that would not have taken extra time. Going forward, standing law reform bodies like the ALRC may even wish to develop a standard format that fits their institutional needs. For example, Lagisz, Samarasinghe, and Nakagawa recently recommended and tested a transparent rapid systematic review format for synthesising built environment research (a field that studies and designs spaces for human activities, relevant for decision- and policymakers).[148] This rapid format contrasts with those used by the four rapid reviews in our study because it requires full transparency.

Alternatively (or in addition), post hoc evaluation tools (sometimes called checklists and assessment tools) can be used to assess whether synthesists achieved a sufficiently high level of transparency.[149] This evaluation can then be amended to the report as a signal about the level of transparency and rigour associated with the report's findings. Similar public signals (ie, badges) have been successful at encouraging transparent research practices in psychology.[150] The data can also be used to measure whether the transparency of law reform synthesis is improving over time.

Second, law reform bodies should adapt their processes to incorporate preregistration of a full protocol as a means of safeguarding primary research and synthesis against bias.[151] In other words, they should make synthesis and data collection protocols public *prior* to that research being conducted.[152] In fact, preregistration can be seen as a mere incremental step from the already iterative and public process by which law reform works. Australian law reform bodies currently work on a model by which public consultations and a public consultation paper (Figure 3) are embedded into the process. As an important tweak to that process, research and synthesis protocols should also be made public, and this should be done before researchers and law reformers see the data (be that primary research or existing articles and studies). This may help in both discouraging changes to the data collection in response to what is found and by making any such change open to discussion and justification (ie, transparent). Depending on the project, a

146 Shinichi Nakagawa and Malgorzata Lagisz, 'How Good Does Our Map of Knowledge Have to Be: A Comment on Berger-Tal et al' (2019) 30(1) *Behavioral Ecology* 13, 13. <https://doi.org/10.1093/beheco/ary137>.

147 Chin, Lagisz and Nakagawa, 'Raw Data' (n 139).

148 Malgorzata Lagisz, Gihan Samarasinghe and Shinichi Nakagawa, 'Rapid Reviews for the Built Environment: Methodology and Guidelines' (Milestone Report Activity, CRC for Low Carbon Living, October 2018).

149 Shea et al (n 82); Woodcock, Pullin and Kaiser (n 82).

150 Mallory C Kidwell et al (n 23). Although note that such signals were less successful in journals from a field that were not already in the midst of a reform movement (suggesting some knowledge of the reasons why open practices are important may be necessary for badges to work): Rowhani-Farid and Barnett (n 23).

151 For a model of preregistration in litigation, see Beerdsen (n 56). In forensic science, see Chin, Growns and Mellor (n 56).

152 Researchers should also be free to publish these protocols, as is becoming common in several fields.

protocol may also state whether any special emphasis will be placed on a certain type of evidence, to avoid situations like we described with the QLRC's report into sexual assault laws.

Another benefit of developing a protocol prior to seeing the data is that this allows for peer review of that protocol when it can actually improve the research plan.[153] Once again, this is not a far cry from the current iterative process of law reform whereby a consultation paper sparks discussions and submissions that are ultimately woven into a final report. Rather, peer review of research plans adds an additional feedback mechanism that meaningfully affects and likely improves the law reform body's research process. For instance, pre-data collection peer review may have headed off the well-founded criticisms of the *RCIRCSA* mock jury study.

Third, and in line with the epigraph that began this Part, we recommend more diversity at law reform bodies. Although we have not systematically studied the backgrounds, including the education, culture, and gender of those involved in law reform, it is common to see royal commissions led by retired judges, who are a relatively homogenous and privileged group. As Oreskes argues, however, diversity improves the research process by exposing existing ways of doing things to critical appraisal from perspectives that do not share the same biases.[154] Much as science itself can benefit from diversity of perspectives,[155] the research conducted by and for law reform should incorporate critical appraisal from individuals that are representative of the diverse communities being served.

When it comes to methodological diversity, Nakagawa and colleagues have suggested that interdisciplinarity – bringing together methodologists and subject-matter experts – drives reliable and efficient synthesis.[156] Greater interdisciplinarity would also assist in both challenging the orthodoxy[157] and providing law reform with in-house technical talent in data management and visualisation. In particular, we suggest engaging with individuals who specialise or co-specialise in the social, biological or information sciences, research methods, and metaresearch.

The request for greater diversity also applies to peer review of protocols by experts in research methodology (rather than eminent figures in the field itself). This converges with emerging practices at leading journals. For instance, *Nature* often engages peer reviewers who specialise in methodology: 'We also routinely require that every meta-analysis or synthesis article submitted to the journal is assessed by at least one expert with demonstrable background in evidence synthesis.'[158] It happens that Australia is home to many metaresearchers,[159] and so Australian

153 Soderberg et al (n 50) 994.

154 Oreskes (n 144).

155 See Neil A Lewis Jr and Jonathan Wai, 'Communicating What We Know, and What Isn't So: Science Communication in Psychology' (2021) 16(6) *Perspectives on Psychological Science* 1242, 1247 <https://doi.org/10.1177/1745691610964062>.

156 Nakagawa et al (n 7) 498.

157 And note that greater diversity can also be achieved by allowing stakeholders to review research protocols before they are executed.

158 'Synthesis Revolution' (n 5) 489.

159 Australian or Australia-based metaresearch groups and organisations include: The MetaMelb research group <https://www.metamelb.org/> at the University of Melbourne; The Deakin Lab for the Meta-

reform bodies are especially well placed to engage individuals with backgrounds that will help them strengthen their methods.

Fourth, and more ambitiously, law reform bodies may begin to include in their preregistrations statements about how they will assess the quality of the evidence they are relying on. Published research is not of uniform quality. And, as we saw with the QLRC controversy (ie, its reliance on one survey of jurors over other studies), law reform bodies sometimes have to make decisions about what research they will place more reliance on. This is a thorny issue[160] because there are few clear-cut indicia what makes some research more reliable than other research (although this is currently under investigation).[161] However, greater transparency about how commissioners will ultimately weigh up the existing body of research can help by both foregrounding that discussion such that it is not affected by what is ultimately found (eg, determining ahead of time that studies with more representative samples will be preferred over others) and by making those decisions more transparent. Moreover, there are some relatively uncontroversial aspects of research that indicate its quality, which can be introduced into registrations when appropriate.[162] For instance, independent replication is a good indicator of its reliability.[163]

Finally, law reform bodies should consider setting up what are known as 'living systematic reviews',[164] as Cochrane has done with currently six living systematic reviews running. Living systematic reviews, as the name suggests, continuously update evidence as new studies come in. However, such reviews require considerable infrastructure as updates are done utilising social media, manual search, and machine-learning based search. Still, this approach fits very well with law reform efforts that seek to have an enduring impact, such as the *RCIRCSA*.[165] Such a mission would be well served by setting up a transparent living review website that collects research conducted in the years after the commission itself.

Analysis of Research (DeLMAR) <https://www.deakin.edu.au/business/research/delmar>; The Inter-Disciplinary Ecology and Evolution Lab <http://www.i-deel.org/> at the University of New South Wales; The Association for Interdisciplinary Meta-research and Open Science ('AIMOS') <https://aimos.community/> which currently has over 250 members; The Australian Reproducibility Network <https://instituteebh.wixsite.com/website-4>.

160 Peter Herbison, Jean Hay-Smith and William J Gillespie, 'Adjustment of Meta-Analyses on the Basis of Quality Scores Should Be Abandoned' (2006) 59(12) *Journal of Clinical Epidemiology* 1249 <https://doi.org/10.1016/j.jclinepi.2006.03.008>.

161 Hannah Fraser et al, 'Predicting Reliability through Structured Expert Elicitation with RepliCATS (Collaborative Assessments for Trustworthy Science)' (Research Paper, University of Melbourne, 17 February 2021).

162 Malcom R Macleod et al, 'Risk of Bias in Reports of in Vivo Research: A Focus for Improvement' (2015) 13(10) *PLoS Biology* e1002273:1–12 <https://doi.org/10.1371/journal.pbio.1002273>.

163 Brian A Nosek and Timothy M Errington, 'What Is Replication?' (2020) 18(3) *PLoS Biology* e3000691:1–8 <https://doi.org/10.1371/journal.pbio.3000691>.

164 Julian H Elliot et al, 'Living Systematic Reviews: An Emerging Opportunity to Narrow the Evidence-Practice Gap' (2014) 11(2) *PloS Medicine* e1001603:1–6 <https://doi.org/10.1371/journal.pmed.1001603>.

165 It detailed its plan for long-term impact in a report: *RCIRCSA Final Report* (n 2) vol 17.

V LIMITATIONS, FURTHER DIRECTIONS, AND CONCLUSIONS

To review, we have canvassed the possibility of deploying open primary research and synthesis practices in the law reform context. This includes preregistration and peer review of synthesis protocols, use of reporting guidelines, open data, and several other practices aimed at improving the transparency of law reform. Many benefits flow from these potential changes to the law reform process, such as increased reliability through managing researcher bias, and increased efficiency by the reports being updateable as new research and data become available. On a more fundamental level, open and transparent research and synthesis are deeply aligned with the principles of law reform bodies, which include the democratisation of their work through transparency, accountability, and open access.[166]

Given these benefits, it is natural to ask why law reform bodies have not taken up these transparent practices. In other words, what hurdles seem to be preventing open law reform? Above, we suggested a lack of interdisciplinarity and divergent research norms between law and other fields may be causing stagnation. As a result, outreach may be needed to explain the benefits of transparent research practices in order to convince law reform bodies to begin to work with research methodology specialists. In this respect, we hope this article provokes an important conversation about open law reform.

More generally, there will be some start-up costs in bringing open and transparent research practices to law reform.[167] These include the expense of training and learning to work with unfamiliar technologies. Moreover, there are likely few individuals with both legal training and research methodology training. This may increase cost because it would require involving two people rather than one. Still, this may be minimised by hiring part-time research help with the required methodological expertise for projects that engage with a great deal of data and evidence. After all, we did find that research specialists seemed to help with the *RCIRCSA* reports. Law reform bodies should also consider developing systems that legal researchers can use with little guidance from methodological experts and introducing training in research methodology (beyond the legal research training staffers receive as part of their educations and on the job).

Start-up costs will likely pay off in the long run. Open data in the context of primary research accelerates discovery because researchers can use open datasets to address new questions. Open synthesis is also valuable because syntheses are not stuck in time; they are capable of being updated as new research is conducted. Furthermore, as we have noted, more efficient formats exist that still prioritise transparency, especially when accompanied by the requirement to follow reporting guidelines and submit checklists. Productive collaborations between

166 See above nn 3–4.

167 For the barriers to transparent primary research, see Christopher Allen and David MA Mehler, 'Open Science Challenges, Benefits and Tips in Early Career and Beyond' (2019) 17(5) *PloS Biology* e3000246:1–14 <https://doi.org/10.1371/journal.pbio.3000246>. For the barriers to open synthesis, see Haddaway et al, 'Open Synthesis' (n 7) 5–6.

industry and metaresearchers have occurred and are capable of delivering transparent and robust evidence in a timely manner.[168] Law reform bodies can follow these models.

Our final suggestions are directed at metaresearchers. While metaresearchers sometimes remark that research-on-research is relevant to law and policy,[169] there appears to be a lack of such work,[170] especially in the context of law reform. Moreover, as noted, law reform may benefit from greater engagement with outside researchers, especially metaresearchers. Therefore, we encourage further metaresearch on law reform. This work can build off established protocols and systems developed in other areas. For instance, health researchers have developed interview and survey methods to better understand how Australian policymakers use research and what they see as the barriers to using research.[171] A similar project in the UK surveyed lawmakers themselves.[172] A project drawing from these approaches would help build upon this article's findings by systematically interviewing and working with law reform staff to help them use and convey research more effectively, and more in line with their stated goals. In short, metaresearch has untapped potential to advance justice.

168 Lagisz, Samarasinghe and Nakagawa (n 148).

169 Miguel et al (n 2) 30; Nakagawa et al (n 7) 498.

170 But see Buckley et al (n 51); de la Guardia, Grant and Miguel (n 51); Mayo-Wilson (n 51); Zinta Zommers et al, 'Burning Embers: Towards More Transparent and Robust Climate-Change Risk Assessments' (2020) 1(10) *Nature Reviews Earth and Environment* 516 <https://doi.org/10.1038/s43017-020-0088-0>.

171 See Sue E Brennan, 'Development and Validation of SEER (Seeking, Engaging with and Evaluating Research): A Measure of Policymakers' Capacity to Engage with and Use Research' (2017) 15(1) *Health Research Policy and Systems* 1 <https://doi.org/10.1186/s12961-016-0162-8>.

172 David Christian Rose et al, 'Improving the Use of Evidence in Legislatures: The Case of the UK Parliament' (2020) 16(4) *Evidence and Policy* 619 <https://doi.org/10.1332/174426420X15828100394351>.