2020

Improving the credibility of empirical legal research: practical suggestions for researchers, journals, and law schools

Kathryn Zeiler
Boston University School of Law

Jason Chin

Alexander DeHaven

Tobias Heycke

Alexander Holcombe

See next page for additional authors

Authors

Kathryn Zeiler, Jason Chin, Alexander DeHaven, Tobias Heycke, Alexander Holcombe, David Mellor, Justin Pickett, Crystal Steltenpohl, and Simine Vazire

IMPROVING THE CREDIBILITY OF EMPIRICAL LEGAL RESEARCH: PRACTICAL SUGGESTIONS FOR RESEARCHERS, JOURNALS, AND LAW SCHOOLS

Jason M. Chin
University of Sydney School of Law

Alexander C. DeHaven
Center For Open Science

Tobias Heycke
GESIS – Leibniz Institute for the Social Sciences

Alexander O. Holcombe
University of Sydney School of Psychology

David T. Mellor

Center for Open Science

Justin T. Pickett
University at Albany, SUNY

Crystal N. Steltenpohl
University of Southern Indiana

Simine Vazire
University of Melbourne School of Psychological Sciences

Kathryn Zeiler
Boston University School of Law

Improving the credibility of empirical legal research: practical suggestions for researchers, journals, and law schools

Jason M. Chin,[1, A] Alexander C. DeHaven,[2, A] Tobias Heycke,[3, A] Alexander O. Holcombe,[4, A] David T. Mellor,[2, A] Justin T. Pickett,[5, A] Crystal N. Steltenpohl,[6, A] Simine Vazire,[7, A] Kathryn Zeiler[8, A]

[1]University of Sydney School of Law, jason.chin@sydney.edu.au

[2]Center For Open Science, david@cos.io, alex.dehaven@cos.io

[3]GESIS – Leibniz Institute for the Social Sciences, tobias.heycke@gmail.com

[4]University of Sydney School of Psychology, alex.holcombe@sydney.edu.au

[5]University at Albany, SUNY, jpickett@albany.edu

[6]University of Southern Indiana, cnsteltenp@usi.edu

[7]University of Melbourne School of Psychological Sciences, simine@gmail.com

[8]Boston University School of Law, kzeiler@bu.edu

[A]Authors listed in alphabetical order

CRediT Statement

Conceptualization: Jason M. Chin.

Data Curation: David T. Mellor.

Investigation: Jason M. Chin.

Project Administration: Jason M. Chin.

Writing - Original Draft Preparation: Jason M. Chin, Crystal N. Steltenpohl, David T. Mellor, Alexander C. DeHaven and Justin T. Pickett.

Writing - Review & Editing: Jason M. Chin, Crystal N. Steltenpohl, David T. Mellor, Alexander C. DeHaven, Alexander O. Holcombe, Justin T. Pickett, Simine Vazire, Kathryn Zeiler and Tobias Heycke.

<u>Abstract</u>

Fields closely related to empirical legal research are enhancing their methods to improve the credibility of their findings. This includes making data, analysis code, and other materials openly available, and preregistering studies. Empirical legal research appears to be lagging behind other fields. This may be due, in part, to a lack of meta-research and guidance on empirical legal studies. The authors seek to fill that gap by evaluating some indicators of credibility in empirical legal research, including a review of guidelines at legal journals. They then provide both general recommendations for researchers, and more specific recommendations aimed at three commonly used empirical legal methods: case law analysis, surveys, and qualitative studies. They end with suggestions for policies and incentive systems that may be implemented by journals and law schools.

Part I. Introduction

Openness and transparency are central to the operation of many legal systems.[1] These virtues are expressed through mechanisms like opening courtrooms to the public and media, and publishing legal decisions.[2] The idea is that by inviting scrutiny, an open legal system is more likely to earn its credibility.[3] Despite the importance of openness and transparency in law, much empirical legal research is still conducted using opaque research practices.[4] This opacity makes it difficult for others to verify and build upon existing findings, and therefore threatens the long-term credibility of the field. In this article, we will provide concrete guidance, based on our own experiences and fields of expertise,[5] that researchers and institutions may use to improve the credibility of empirical legal research (ELR).[6]

Several fields – some closely related to ELR, and some that lend methodologies to ELR – are undergoing a "credibility revolution".[7] In the context of empirical research, "credibility" is

---

[1] "Publicity is the very soul of justice […] in comparison of publicity, all other checks are of small account. It is to publicity, more than to everything else put together, that the English system of procedure owes its being the least bad system as yet extant, instead of being the worst." JEREMY BENTHAM, WORKS OF JEREMY BENTHAM VOL. 4, (Bowring, ed., 1843) at 305, 316-7; *Re Vancouver Sun* [2004] 2 SCR 332; *In the Matter of an Application by Chief Commissioner of Police (Vic)* (2005) 214 ALR 422.

[2] See Mitchell T. Galloway, *The States Have Spoken: Allow Expanded Media Coverage of the Federal Courts*, 21 VAND. J. ENT. & TECH. L. 77 (2018).

[3] Tom R. Tyler, *Procedural justice and the courts*, 44(1-2) COURT REV. 26 (2007); "Sunlight is ... the best of disinfectants." *Buckley v. Valeo*, 424 U.S. 1, 67 (1976) (per curiam) (quoting LOUIS D. BRANDEIS, OTHER PEOPLE'S MONEY AND HOW THE BANKERS USE IT 62 (1933)); Bentham, *supra* note 1.

[4] Lee Epstein & Gary King, *The Rules of Inference*, 69 U. Chi. L. Rev. 1 (2002); Kathryn Zeiler, *The Future of Empirical Legal Scholarship: Where Might We Go from Here?* 66 J. LEGAL EDUC. 78 (2016); Hillel J. Bavli, *Credibility in Empirical Legal Analysis*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3434095 (accessed 2020).

[5] These are: analysis of case law, survey methods, qualitative methods, changing research cultures.

[6] For the purposes of this article, we will adopt a broad definition of ELR as "…any attempt to collect and analyze a set of data for more than anecdotal purposes…" Russell Korobkin, *Empirical Scholarship in Contract Law: Possibilities and Pitfalls*, 2002(4) U. ILL. L. REV. 1033 (2002) at 1055; see also Shari Seidman Diamond, *Empirical Marine Life in Legal Waters: Clams, Dolphins, and Plankton*, 2002(4) U. ILL. L. REV. 803 (2002).

[7] Simine Vazire, *Implications of the Credibility Revolution for Productivity, Creativity, and Progress*, 13(4) PERSP. PSYCHOL. SCI. 411 (2018); Joshua D. Angrist and Jörn-Steffen Pischke, *The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con out of Econometrics*, 24(2) J. ECON. PERSP. 3 (2010); Marcus R. Munafò et al., *A manifesto for reproducible science*, 1 NAT. HUM. BEHAV. 1 (2017) at 1, 5. See also NATIONAL ACADEMIES OF SCIENCES, ENGINEERING, AND MEDICINE, OPEN SCIENCE BY DESIGN: REALIZING A VISION FOR 21ST CENTURY RESEARCH (2018) at 107 ("NASEM Report").

used to mean that a study's methodology and results are reported transparently so that they can be verified and repeated by others, that errors are caught and corrected, and that the conclusions are updated and well-calibrated to the strength of the evidence. In other words, credibility does not mean that findings are perfectly accurate or error-free, but that researchers prioritize transparency and calibration, such that errors are easy to catch and correct. A credible field is one where the findings are reported with appropriate levels of confidence and certainty and errors are likely to get corrected, such that when a finding withstands scrutiny and becomes well-established, it is very likely to be true.[8]

Credible research presents many additional advantages: (1) it is "reproducible", meaning that its data, methods, and analysis are reported with enough detail such that other researchers can verify conclusions and correct them if needed;[9] (2) it is more efficient because reproducible workflows allow other researchers to build upon existing work and to test new questions;[10] (3) its findings are "replicable", meaning they can be confirmed through testing with new data;[11] (4) its claims are well-calibrated such that bodies that fund and rely on this research can employ them with more confidence,[12] and; (5) it inspires greater public trust.[13] Many of these benefits were recently encapsulated in a 2018 National Academy of Sciences, Engineering, and Medicine

---

[8] Munafò et al., *id.* at 5: "Claims become credible by the community reviewing, critiquing, extending and reproducing the supporting evidence. However, without transparency, claims only achieve credibility based on trust in the confidence or authority of the originator. Transparency is superior to trust."; Simine Vazire & Alex O. Holcombe, *Where Are The Self-Correcting Mechanisms in Science?*, https://psyarxiv.com/kgqzt/.

[9] Shareen A. Iqbal et al., *Reproducible Research Practices and Transparency across the Biomedical Literature*, 14(1) PLoS. BIOL. e1002333 (2016); Tom E. Hardwicke et al., *Calibrating the Scientific Ecosystem Through Meta-Research*, 7 ANNU. REV. STAT. APPL. 11 (2020) at 16.

[10] Timothy H. Vines et al., *The Availability of Research Data Declines Rapidly with Article Age*, 24 CURR. BIOL. 94 (2014); Iain Chalmers & Paul Glasziou, *Avoidable waste in the production and reporting of research evidence*, 374 LANCET 86 (2009).

[11] Brian A. Nosek & Timothy M. Errington, *Making sense of Replications*, 6 ELIFE e23383 (2017).

[12] Hardwicke et al., *Calibrating the Scientific Ecosystem*, *supra* note 9.

[13] CARY FUNK ET AL., *Trust and Mistrust in Americans' Views of Scientific Experts*, Pew Research Center (2019) 24; Simine Vazire, *Quality Uncertainty Erodes Trust in Science*, 3(1) COLLABRA: PSYCHOLOGY 1 (2017).

("NASEM") consensus report about openness and transparency in science, which said: "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."[14]

These advantages are even greater in fields like ELR whose impact frequently extends beyond academia to real world issues.[15] For example, ELR is regularly cited by courts in the United States.[16] It is also relied on by policy-makers, and sometimes commissioned and funded by law reform bodies.[17] In addition, empirical legal research is often publicly funded (like many fields), so the public has an interest in seeing its methods and data be publicly accessible.[18]

In this article, our primary objective is to provide specific steps legal researchers and institutions can take to enhance research credibility.[19] Prior to doing that, Part II will review and contextualize existing concerns about ELR. In that part, we will provide a novel review of law journal guidelines to gauge the extent to which they promote credible practices. This analysis leads into our recommendations for concrete steps that researchers, journals, and institutions can take. In Part III, we will discuss three common empirical methodologies in ELR and how they

---

[14] NASEM, *supra* note 7 at 107.

[15] For a similar argument in medicine, see Bob Carlson, Putting Oncology Patients at Risk, 9(3) BIOTECHNOL. HEALTHC. 17 (2012); in forensic science, see Jason M. Chin, Gianni Ribeiro & Alicia Rairden, *Open Forensic Science*, 6(1) J. LAW BIOSCI. 255 (2019).

[16] Zeiler, *supra* note 4 at fn 34; Lee Petherbridge & David L. Schwartz, An Empirical Assessment of the Supreme Court's Use of Legal Scholarship, 106(3) NW. U. L. REV. 995 (2012).

[17] See David Hamer, *The Significant Probative Value of Tendency Evidence*, 42 MELB. U. L. REV. 506 (2019) at fn. 52; Christopher Uggen & Michelle Inderbitzin, *Public Criminologies*, 9(4) CRIMINOL. & PUB. POL. 725 (2010).

[18] Michael Gibbons, *Science's new social contract with society*, 402 NATURE C81 (1999). See also Todd J. Vision, *Open Data and the Social Contract of Scientific Publishing*, 60(5) BIOSCI. 330 (2010).

[19] Indeed, one established barrier to data sharing is lack of knowledge of how to do it, see Laure Perrier, Erik Blondal & Heather MacDonald, *The views, perspectives, and experiences of academic researchers with data sharing and reuse: A meta-synthesis*, 15(2) PLOS ONE (2020) at 13: "For many disciplines, data sharing was a new activity that was typically imposed by funding agencies or journals. As a result, researchers were looking for services or resources that would help with this task." For a practical guide for credible psychology research, see Olivier Klein et al., *A Practical Guide for Transparency in Psychological Science*, 4 COLLABRA: PSYCHOLOGY 4 (2018).

may be conducted more credibly. Then, in Part IV, we turn to journals and societies, and the steps they may take to promote research credibility. Part V provides some final reflections on the path forward.

Part II. The credibility of empirical legal research

We will now contextualize ELR within the broader movement afoot in social science towards increased credibility. We will start by reviewing reforms that are becoming mainstream in social science and the concerns that inspired those reforms. We will then discuss the challenges particular to ELR, ending with an analysis of law journal publication guidelines, finding that there is considerable heterogeneity and room for improvement.

1. The credibility revolution in social science

Social scientific fields akin to ELR have recently taken steps to enhance their credibility.[20] Consider, for instance, the 2018 State of Social Science (3S) survey, which asked hundreds of researchers in psychology, economics, sociology and political science about whether and when they had first made public their data and materials (e.g., instruments like surveys, images or sound files presented to participants), and whether they had preregistered a study (i.e., publicly registered their hypotheses and methods before running the study, see below).[21] The survey showed considerable increases in all three self-reported behaviors over the past several years (Figure 1).

[**Figure 1 about here**]

---

[20] See *infra* at Part II.1(a)-(c).

[21] Garret Christensen et al., *Open Science Practices are on the Rise: The State of Social Science (3S) Survey*, https://osf.io/preprints/metaarxiv/5rksu/ (accessed 2020). Figure 1 is reproduced under a CC-By Attribution 4.0 International license.

In psychology and economics, these reforms were inspired, in part, by the surprising failures to replicate findings published in prestigious journals.[22] These failures to replicate were difficult to ignore because of the high methodological quality of the replication studies, often pre-registered and in some cases involving sample sizes several times larger than the originals and with data collected across multiple labs. The results of these replication projects likely contributed to the opinions expressed by researchers in a survey in which 52% of respondents said there was a significant crisis in science, 38% said there was a slight crisis, and only 3% said there was no crisis.[23]

The causes of this perceived crisis can be broken down into those that are preventable and those that are more difficult to control. As to the preventable (which we are most interested in),[24] self-report surveys of researchers are documenting widespread use of questionable research practices (QRPs) in many fields (e.g., psychology, economics, evolutionary biology).[25] QRPs take advantage of undisclosed flexibility in the research process to allow researchers to make their findings seem cleaner and more persuasive than they actually are (e.g., one widespread QRP is making the decision to exclude an outlier after seeing the data). Whether conscious or

---

[22] Open Science Collaboration (OSC), *Estimating the Reproducibility of Psychological Science*, 349 SCIENCE 3451 (2015); Richard A. Klein et al., *Investigating variation in replicability: A 'many labs' replication project*, 45(3) SOC. PSYCHOL. 142 (2014); Richard A. Klein et al., *Many Labs 2: Investigating Variation in Replicability Across Samples and Settings*, 1 ADV. METH. & PRACT. PSYCHOL. SCI. 443 (2018); Charles Ebersole et al., *Many Labs 3: Evaluating participant pool quality across the academic semester via replication*, 67 J. EXP. SOC. PSYCHOL. 68 (2016); Colin F. Camerer et al., *Evaluating replicability of laboratory experiments in economics*, 351 SCIENCE 1433 (2016); Colin F. Camerer et al., *Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015*, 2 NAT. HUM. BEHAV. 637 (2018).

[23] Monya Baker, *1,500 scientists lift the lid on reproducibility*, 533 NATURE 452 (2016); See also John P. A. Ioannidis, *Why Most Published Research Findings Are False*, 8 PLOS MED. E124 (2005).

[24] Unpreventable causes include the inherent noise in any data collection and the possibility of unmeasured variables driving the originally discovered effect (i.e., hidden moderators). For instance, when studying human social processes, the study's setting and context may be difficult or impossible to recreate, and so researchers may find different results upon replication: see Klein et al., *Many Labs 2*, *supra* note 22 at 482.

[25] Leslie K. John et al., *Measuring the Prevalence of Questionable Research Practices With Incentives for Truth Telling*, 23(5) PSYCHOL. SCI. 524 (2012); Sarah Necker, *Scientific misbehavior in economics*, 43 RES. POL. 1747 (2014); Hannah Fraser et al., *Questionable research practices in ecology and evolution*, 13(7) PLOS ONE 1 (2018).

not, these practices make results seem more surprising, and therefore publishable, but diminish the credibility of results. For instance, data-contingent exclusion of outliers demonstrably inflate a field's false positive rate.[26]

Another preventable cause is publication bias, which is the tendency for only statistically significant findings to be published.[27] This can make the available literature a poor gauge of the actual strength of a finding.[28]

Even among published studies, most do not make their materials and data available.[29] This prevents mistakes from being corrected by other researchers.[30]

As we noted, several fields are increasingly adopting reforms that respond to these controllable sources of error.[31] We will now briefly review some of those reforms, which we will revisit in greater detail when we discuss how they can be leveraged by empirical legal researchers in Part III. Note that early results suggest these reforms can be effective. For instance, whereas the false discovery rate for traditional psychological studies hovers around 50%, a 2020 study found that studies using several of the below reforms (e.g., preregistration,

---

[26] Joseph P. Simmons et al., *False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant*, 22(11) PSYCHOL. SCI. 1359 (2011).

[27] Annie Franco, Neil Malhotra & Gabor Simonovits, *Publication bias in the social sciences: Unlocking the file drawer*, 345(6203) SCIENCE 1502; Daniele Fanelli, *Negative results are disappearing from most disciplines and countries*, 90 SCIENTOMETRICS 891 (2012); Ioana Allina Cristea & John P. A. Ioannidis, P values in display items are ubiquitous and almost invariably significant: A survey of top science journals, 13(5) PLOS ONE e0197440 (2018).

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

[29] Tom E. Hardwicke et al., *An empirical assessment of transparency and reproducibility-related research practices in the social sciences (2014–2017)*, 7 R. SOC. OPEN SCI. (2020).

[30] For an example of the literature self-correcting when data is open, see Joscha Legewie, *Retraction of the Research Article: "Police Violence and the Health of Black Infants"*, 5(12) SCI. ADV. eaba5491 (2019). For a review of all preventable causes of failures to replicate, see Hardwicke et al., *Calibrating the Scientific Ecosystem*, *supra* note 9.

[31] For uncontrollable sources, see *supra* note 24.

methods transparency) were replicated at the rate that would be expected under traditional

statistical assumption.[32]

(a) Preregistration

Preregistration, also known as prospective study registration or a pre-analysis plan,

became required by law in some jurisdictions for clinical medical research due, in part, to

widespread concerns about the public health implications of publication bias (i.e., non-

publication of entire studies unfavorable to the drug manufacturer, partial reporting of results,

and outcome switching).[33] Preregistration involves, prior to data collection, submitting the

hypotheses, methods, data collection plan, and analysis plan to a common registry.[34]

Preregistration makes the existence of unreported studies more findable (e.g., for meta-

analyses)[35] and can discourage (or at least make detectable) QRP usage by preserving a record of

the methodology as it was before the data were viewed.[36] They are associated with reduced

publication bias.[37] Researchers following a preregistered analysis plan may also be less likely to

mistake prediction with postdiction. In this way, preregistration can play a similar role to results-

blind analysis, which was developed in physics but is applicable to social science.[38] In other

words, they may be less likely to present an exploratory finding (e.g., through mining the data for

---

[32] John Protzko et al., *High Replicability of Newly-Discovered Social-behavioral Findings is Achievable*, https://psyarxiv.com/n2a9x (accessed 2020).

[33] Charles Piller, *Transparency on trial*, 367(6475) SCIENCE 240 (2020); Kay Dickersin & Iain Chalmers, *Recognizing, investigating and dealing with incomplete and biased reporting of clinical research: from Francis Bacon to the WHO*, 104 J. R. SOC. MED. 104 (2011) 532.

[34] Brian A. Nosek et al., *The Preregistration Revolution*, 115(11) PNAS 2600 (2018).

[35] This is because registries of studies, many of which may not yet be published, can be searched, see OSF Registries, https://osf.io/registries (accessed 2020).

[36] John et al., *supra* note 25.

[37] Robert M. Kaplan & Veronica L. Irvin, *Likelihood of Null Effects of Large NHLBI Clinical Trials Has Increased over Time*, 10(8) PLoS ONE e0132382 (2015).

[38] Robert MacCoun and Saul Perlmutter, *Hide results to seek the truth*, 526 NAT. 187 (2015); For the postdiction-prediction problem, see Nosek et al., *The Preregistration Revolution*, *supra* note 34.

statistical significance) as one they predicted. If detailed pre-registered plans are followed, this

preserves the statistical validity of analyses with respect to error control.[39]

Although many fields are embracing preregistration, it should not be seen as a panacea

for all threats to research credibility. Rather, it is an important tool that can be employed with

other reforms aimed at shifting incentives towards getting it right (versus getting it published).

Registered reports are another excellent example of such a shift.

Registered reports ("RRs") are a new type of article where the peer review process is

restructured such that reviewers evaluate the proposed methods and justification for the study

before the study has been conducted and the results are known.[40] If the editor accepts the

proposal, it is then guaranteed to be published if the authors follow through on that plan, which is

then preregistered. Publications are, therefore, not selected based on results but the research

question and method. Like preregistration, RRs can reduce publication bias and QRPs. They can

also result in improved methodology as reviewers provide criticism and advice regarding

methods before data is collected. Over 260 journals now accept registered reports, a number that

has ballooned from under 5 in 2014.[41]

[**Figure 2 about here**]

Recent analyses find registered reports are more likely to report null results (see Figure

2).[42] This is salutary because the proportion of published literature that contains positive results

is so high (~95%) that it is almost guaranteed that many of the positive results are false

---

[39] Eric-Jan Wagenmakers, et al., *An Agenda for Purely Confirmatory Research*, 7(6) PERSP. PSYCHOL. SCI. 632 (2012).

[40] Chris Chambers, *What's next for registered reports?*, 573 NATURE 187 (2019).

[41] *Id.*

[42] Anne M. Scheel, Mitchell Schijen & Daniël Lakens, *An excess of positive results: Comparing the standard Psychology literature with Registered Reports*, https://psyarxiv.com/p6e9c (accessed 2020). Figure 2 is reproduced under a CC-By Attribution 4.0 International license; see also Christopher Allen & David M. A. Mehler, *Open science challenges, benefits and tips in early career and beyond*, 17(5) PLOS BIOL. e3000246.

positives.[43] The rate of positive results in the (small) Registered Reports literature (about 45% according to two studies) is more realistic and consistent with the rate of positive results in large-scale direct replication projects in the social sciences.[44]

(b) Open data, code, and materials

Journals and researchers are increasingly embracing open data and code.[45] For example, several economics and political science journals now require that authors make their data public (e.g., uploaded to a public repository) and provide the code that translates their data into the stated results (see Part II.1(b) below).[46] Similarly, researchers in many fields increasingly share their study materials and many journals encourage or require this practice.[47]

(c) Badges, checklists, and other reforms

Article badges for practices like preregistration, open materials, and open data present a way to recognize and reward open practices without requiring them.[48] One analysis found that after a journal adopted a badge policy, published articles had much higher rates of practices associated with those badges.[49] That said, the effect appears to be stronger if actors in the field are already aware of the benefits of such practices.[50]

---

[43] Fanelli, *supra* note 27; Cristea & Ioannidis, *supra* note 27.

[44] See sources at *supra* note 22.

[45] Christensen et al., *supra* note 21; Center for Open Science, *TOP Guidelines*, https://cos.io/top/ (accessed 2020); Mark D. Wilkinson et al., *The FAIR Guiding Principles for scientific data management and stewardship*, 3 SCI. DATA 160018 (2016).

[46] Christina Blanco-Perez & Abel Brodeur, *Transparency in empirical economic research*, https://wol.iza.org/articles/transparency-in-empirical-economic-research/long (accessed 2020). Political science journals often verify that the analytic produces the results presented in the article, see American Journal of Political Science, *AJPS Verification Policy*, https://ajps.org/ajps-verification-policy/ (accessed 2020)

[47] Christensen et al., *supra* note 21; Center for Open Science, *TOP Guidelines*, *supra* note 45.

[48] Center for Open Science, *Open Science Badges*, https://cos.io/our-services/open-science-badges/ (accessed 2020).

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

[50] Anisa Rowhani-Farid & Adrian G. Barnett, *Badges for sharing data and code at Biostatistics: an observational study*, 7(90) F1000RESEARCH (2018).

Many fields have also created checklists that researchers are required or encouraged to submit with their manuscripts (e.g., acknowledging they have reported all produced estimates).[51] Some of these checklists have been associated with fuller reporting of, for example, a study's methodological limitations.[52]

Finally, reforms to the peer review process are spreading.[53] These include: open peer review models in which peer reviews are published along with the articles; continuing peer review (in which commentaries can be appended to existing publications); and changes in peer review criteria, such as judging articles based on credibility instead of novelty.[54] As we will discuss further below, journals are also increasingly adopting guidelines that encourage or require practices like open data and preregistration.[55]

2. Concerns about the credibility of empirical legal research

Expressed concerns about the credibility of ELR actually predate much of the discussion above, but have produced no lasting reforms or initiatives that we could find.[56] In 2001, for instance, Epstein and King reviewed many empirical legal studies and found errors of inference

---

[51] Equator Network, *Enhancing the QUAlity and Transparency Of health Research*, https://www.equator-network.org/ (accessed 2020); Balazs Aczel et al., *A consensus-based transparency checklist*, 4 NAT. HUM. BEHAV. 4 (2020).

[52] 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*, 1 SYST. REV. 60 (2012); SeungHye Han, *A checklist is associated with increased quality of reporting preclinical biomedical research: A systematic review*, 12(9) PLoS ONE e0183591 (2017).

[53] S. P. J. M. (Serge) Horbach & W. (Willem) Halffman, *The changing forms and expectations of peer review*, 3(8) RES. INTEGR. PEER REV. 1 (2018).

[54] *Id.* at 5-9; For a specific example of a new peer review platform, see *Pubpeer*, https://pubpeer.com/ (accessed 2020).

[55] B. A. Nosek et al., *Promoting an open research culture*, 348(6242) SCIENCE 1422 (2015).

[56] See Epstein & King, *supra* note 4; David E. Van Zandt, *Discipline-Based Faculty*, 53 J. LEGAL EDUC. 332 (2003) at 332.

in all of them. In many of these, the errors stemmed from the conclusions not being based on reproducible analyses.[57] Over a decade later, little seems to have changed.[58]

ELR faces many of the same challenges found in other social sciences and so perhaps it is not surprising that questions would be raised about its practices. In many cases, the relationship is direct: ELR practices often borrow from cognate disciplines, like economics and psychology, two fields whose historic practices contributed concerns about replication and reproduction. ELR also typically operates in a research environment similar many others, in which there is an incentive to publish frequently, perhaps at the cost of quality and rigor.[59] In such environments many journals also appear to favor novel and exciting findings, without a concomitant emphasis on methodology. This combination of incentives creates an ecosystem in which low credibility research is rewarded, and those who engage in more rigorous practices are driven out of the field.[60]

But, ELR also faces its own set of challenges. Much research in this field is published in generalist law journals that may rarely receive empirical work. Some of these journals are edited by students, many of whom cannot be expected to have the appropriate background to evaluate empirical methods.[61] Student editors are also likely swayed by the eminence of authors, which

---

[57] Epstein & King, *id.* at 38-45.

[58] Zeiler, *supra* note 4 surveyed many of the same problems over a decade later.

[59] In science generally, see Brian A. Nosek, Jeffrey R. Spies & Matt Motyl, *Scientific Utopia: II. Restructuring Incentives and Practices to Promote Truth Over Publishability*, 7(6) PERSP. PSYCHOL. SCI. 615 (2012); Paul E. Smaldino & Richard McElreath, *The natural selection of bad science* 3 R. SOC. OPEN SCI. 160384. In law, see Zeiler, *supra* note 4 at 79-80, 87-98. Note, however, that some law schools may differ from the broader research ecosystem in that some have typically enjoyed considerable access to research funding, Epstein & King, *supra* note 4 at 115.

[60] Smaldino & McElreath, *id.*

[61] Zeiler, *supra* note 4 at 78-9; Shari Seidman Diamond, *Empirical Legal Scholarship: Observations on Moving Forward*, 113 NW. U. L. REV. 1229 (2019) at 1230-1.

itself is a biasing force in peer review.[62] As a result, they rely more heavily on authors' status, with less emphasis on methodological quality.

Turning to the authors themselves, many empirical legal researchers possess a primarily legal background. As a result, they may not have the specialized statistical and methodological knowledge required to ensure their work is credible (which, as the replication crisis has brought to light, many people with extensive training in social science also lack). Furthermore, as trained advocates, some authors may be culturally inclined towards strong rhetoric that may, at times, not be entirely justified by the data.

3. Are law journals promoting research credibility?

Journals represent an important pressure point on research practices because they choose what to publish and control the form in which research is reported (e.g., by encouraging or requiring that raw data and code be published along with the typical manuscript narrative). Accordingly, it is useful to ask: to what extent are ELR journals promoting credible practices?

As we saw above, many journals in fields outside of law have begun to reform their guidelines to encourage authors to engage in behaviors like posting their data and preregistering their hypotheses. Much of this was spurred by the development of the *Transparency and Openness Promotion (TOP) Guidelines*, and its goal of "promoting an open research culture".[63] The original TOP Guidelines cover 8 standards, each of which can be implemented in one of three levels of increasing rigor. A "0" indicates that the standard does not comply with TOP, for example a policy that merely encourages data sharing, or says nothing at all about data sharing. Levels 1-3 vary by standard, but for data transparency correspond to: 1) Journal requires

---

[62] Simine Vazire, *Our obsession with eminence warps research*, 547 NAT. 7 (2017).
[63] Nosek et al., *Promoting an open research culture*, *supra* note 55.

disclosure of whether or not data are available (e.g. a data availability statement), 2) Journal

requires data sharing as long as it is ethically feasible, 3) Journal computationally verifies that

data can be used to reproduce the results presented in the paper.

The Center for Open Science recently developed the *TOP Factor*,[64] an open database for

evaluating and scoring journal policies against the standards set in the *TOP Guidelines*. The

database was developed to enable communities served by the journals to more readily evaluate a

journal's policies related to open science practices and allow for easy comparison among

journals within a discipline. The TOP Factor is intended to help journals receive credit for having

more open science practices, facilitating further adoption and change towards more open science

norms in the scholarly literature. It is distinct from existing journal metrics, such as the Journal

Impact Factor and Altmetrics, in that it evaluates practices that are directly associated with the

scientific process.[65] Accordingly, the *TOP Factor* is orthogonal to factors like novelty and

surprisingness of study outcomes, measures that are of interest but nonetheless overvalued.[66] The

TOP Factor measures the original 8 TOP guidelines as well as whether the journal has policies to

counter publication bias (e.g., by accepting registered reports) and whether it awards badges.

To determine to what extent law journals are promoting research credibility, we scored

them with the TOP Factor and entered the results into the COS's larger database under

"Empirical Legal Research" (see also Table 1 for a truncated version).[67] We chose the top 25

---

[64] Center for Open Science, *Top Factor*, http://topfactor.org (accessed 2020). For more information, see Center for Open Science, *TOP Guidelines*, *supra* note 45.

[65] Melissa S. Anderson, Brian C. Martinson & Raymond De Vries, *Normative Dissonance in Science: Results from a National Survey of U.S. Scientists*, 2(4) J. EMPIR. RES. HUM. RES. ETHICS 3 (2007).

[66] Roger Giner-Sorolla, *Science or Art? How Aesthetic Standards Grease the Way Through the Publication Bottleneck but Undermine Science*, 7(6) PERSO. PSYCHOL. SCI. 562 (2012).

[67] Center for Open Science, *TOP Factor*, *supra* note 64. The raw data for all journals can be found here: Center For Open Science, *top-factor.csv*, https://osf.io/qatkz/ (accessed 2020). The raw data for just empirical legal journals can be found here: The Authors, *TOP Factor – Law Journals.xls*, https://osf.io/hk3dt/?view_only=c8c39edb8b4a417db87660e01a46b848 (accessed 2020). The rating rubric can be

student-edited law journals in the United States from the *Washington and Lee Law Journal Rankings* (2018). That ranking is based on citations from other U.S. law journals and legal decisions. We also scored 6 peer-reviewed law journals that we know regularly publish empirical research (*Journal of Empirical Legal Studies*, *Journal of Legal Studies*, *Journal of Law and Economics*, *American Law and Economics Review*, *Journal of Legal Analysis, and Law, Probability, and Risk*).

As can be seen in Table 1, our results are consistent with existing concerns about the level of methodological rigor that student-edited law journals require and promote. Only 3 of the 25 student journals we scored received total scores above 0. Those were *Yale Law Journal* (4), *Stanford Law Review* (2), and *NYU Law Review* (2). NYU and Stanford received their 2s for data transparency by requiring posting of data subject to countervailing reasons. For instance, *Stanford Law Review*'s policy is:

"Replicability: At a minimum, empirical works must document and archive all datasets so that third parties may replicate the published findings. These datasets will be published on our website. The Stanford Law Review will make narrow exceptions on a case-by-case basis, particularly if the datasets involve issues of confidentiality and/or privacy."[68]

Yale Law Journal received a 2 for data transparency for having a data policy similar to NYU and Stanford and another 2 for applying that rule to analytic code.

The 6 peer-reviewed journals fared better on the *TOP Factor*, but there is considerable room to improve. The relatively high scores for some of these journals perhaps come from interfacing with economics, a field in which computational verification of reported findings is

---

found here Center for Open Science, *TOP-factor-rubric.docx*, https://osf.io/t2yu5/ (accessed 2020). Results are up to date as of July 1, 2020.

[68] Stanford Law Review, *Our Submissions Review Process*, https://www.stanfordlawreview.org/submissions/article-submissions/ (accessed 2020).

becoming more mainstream.[69] For example, the *Journal of Legal Studies* and *Journal of Law and Economics* have expressly adopted the data and materials guidelines used by economics journals.[70] The *American Law and Economics Review* has policies for data and code, but they are not as demanding. The *Journal of Legal Analysis* has the strongest guidelines for data citation, having adopted the *Future of Research Communications and e-Scholarship (FORCE11) Joint Declaration of Data Citation Principles*.[71] Troublingly, the *Journal of Empirical Legal Studies* and *Law, Probability and Risk* both score 0 overall.

None of the 31 journals we studied have joined the 256 journals in other fields that accept Registered Reports. Further, none award badges, recommend the use of reporting guidelines, or any policies about replication studies.

Part III. Guidance for researchers seeking to improve their research credibility

We will now further elaborate on some of the credibility reforms we discussed in the previous part. First, we will provide some general guidance for empirical legal researchers interested in implementing these reforms. As part of that discussion, we will highlight resources that are particularly appropriate for social scientific research and widely-used guidelines that can be adapted for ELR methodologies. Then, we go on to discuss the application of these reforms to three mainstream empirical legal methodologies: case law analysis, surveys, and qualitative methods. As we will discuss, these general recommendations are all subject to qualifications, such as the ethics of sharing certain types of data.

---

[69] Christensen et al., *supra* note 21.
[70] The Journal of legal Studies, *Data Policy*, https://www.journals.uchicago.edu/journals/jls/data-policy (accessed 2020); The Journal of Law and Economics, *Data Policy*, https://www.journals.uchicago.edu/journals/jle/data-policy (accessed 2020).
[71] Force11, *Joint Declaration of Data Citation Principles – Final*, https://www.force11.org/datacitationprinciples (accessed 2020).

1. General recommendations

(a) Preregister your studies

Empirical legal researchers interested in enhancing their work's transparency can preregister their studies using platforms like the OSF,[72] the American Economic Association registry,[73] or AsPredicted.[74] These user-friendly services create a timestamped, read-only description of the project.[75] The registration can be made public either immediately or be embargoed for a pre-specified amount of time, or indefinitely. When reporting the results of preregistered work, the author should follow a few best practices. First, include a link to the preregistration so that reviewers and readers can confirm what parts of the study were pre-specified. Second, the authors should report the results of all pre-specified analyses, not just those that are most significant or surprising. Third, any unregistered parts of the study should be transparently reported, ideally in a different sub-section of the results. These "exploratory" analyses should be presented as preliminary, testable hypotheses that deserve confirmation. Finally, any changes from the pre-specified plan should be transparently reported. These changes, some of which will be trivial and some of which may substantially alter the interpretability of the results, can be better evaluated if they are clearly described.

Several templates are available to guide researchers through a preregistration process.[76] Some are quite specific, leading the researcher through several questions about their study's background, hypotheses, sampling, and design. On the other end of the spectrum are those that

---

[72] Center for Open Science, *OSF Preregistration*, https://osf.io/prereg/ (accessed 2020).

[73] The American Economics Association, *The American Economic Association's registry for randomized controlled trials*, https://www.socialscienceregistry.org/ (accessed 2020).

[74] University of Pennsylvania Wharton, Credibility Lab, *AsPredicted*, https://aspredicted.org/ (accessed 2020).

[75] See Nosek et al., *The Preregistration Revolution*, *supra* note 34; Wagenmakers et al., *supra* note 39.

[76] Anna Elisabeth van't Veer & Roger Giner-Sorolla, *Pre-registration in social psychology—A discussion and suggested template*, 67 J. EXP. SOC. PSCYHOL. 2 (2016); Center for Open Science, *Templates of OFS Registration forms*, https://osf.io/zab38/wiki/home/ (accessed 2020).

give the researcher free rein to describe the study in as little or as much detail as they like. We are not aware of any templates specifically designed for empirical legal research (although this would be a worthwhile project), but existing templates are well-suited to both experimental and observational research. Below (Part III.2), we will discuss how preregistration might operate in various ELR paradigms.

(b) Open your data and analytic code

Empirical legal researchers who wish to improve their work's reusability and verifiability, and who wish to ensure their efforts are not lost to time (as research ages, underlying data and materials become increasingly unavailable because authors are not reachable),[77] have many options open to them. Many free repositories have been developed to assist with research data storage and sharing. Given the availability of these repositories, the fact that the publishing journal does not host data itself (or make open data a requirement) is not a good reason in itself to refrain from sharing. Rather, authors can simply reference a persistent locator (such as a Digital Object Identifier, DOI) provided by a repository.

In Table 2, we display a selection of repositories that may be of particular interest to empirical legal researchers, along with some key features of those repositories. This is based, in part, on the work of Oliver Klein and colleagues, who explained the characteristics of data repositories that researchers should consider when deciding which service to use.[78] These include: whether the service provides persistent identifiers (e.g., DOIs), whether it enables citation to the data; whether it ensures long-term storage and access to the data; and, whether it complies with relevant legislation.

---

[77] Vines et al., *supra* note 10.
[78] Klein et al., *A practical guide*, *supra* note 19 at 6.

Access to data alone already provides a substantial benefit to future research, but researchers can do more.[79] In this respect, researchers should strive to abide by the Global Open (GO) FAIR Guiding Principles, which were developed by an international group of academics, funders, industry representatives, and publishers (and endorsed by the NASEM in 2018).[80] These principles are that data should be Findable, Accessible, Interoperable, and Reusable (i.e., FAIR). We have already touched on findable (e.g., via a persistent identifier) and accessible data (e.g., via a long-term repository). Interoperable data is data that can be easily combined with other data and used by other systems (e.g., through code explaining what variables mean). And, reusable data typically means data that have a license that allows reuse and "metadata" that fully explains its provenance. One helpful practice to improve interoperability and reusability is to associate data with a "codebook", or file explaining the meaning of variables. The Center for Open Science maintains a guide to interoperability and reusability relevant to social science.[81]

Open and FAIR data is important to the future of ELR. For example, FAIRness allows researchers to leverage multiple datasets to perform meta-analyses and systematic reviews aimed at better understanding the robustness of a finding. It also enables researchers to combine and compare data across jurisdictions to test new hypotheses.

As to what should be shared, we recommend starting from the presumption that all raw data will be shared and then identifying any necessary restrictions and barriers.[82] Those restraints

---

[79] NASEM, *supra* note 7 at 28: "Data that are open and FAIR will maximize the impact of open science."

[80] Wilkinson et al., *supra* note 45; the NASEM Report, Id. at 53, provides a useful example of how FAIR operates: "An example of FAIR data for human use is provided by public webpages. Search engines have made many such pages findable and they are usually either immediately accessible or accessible via a paywall. Since these pages are designed for human readers, they are made (more or less) interoperable by the readers' knowledge of the language and the subject matter. Pages are often reusable by cut-and-paste document editing tools."

[81] Center for Open Science, *Best practices*, https://help.osf.io/hc/en-us/categories/360001530634-Best-Practices (accessed 2020); See also Klein et al., *A practical guide for transparency in psychological science (supplementary material)*, https://psych-transparency-guide.uni-koeln.de/data-documentation.html (accessed 2020).

[82] Klein et al., *A practical guide supra* note 19 at 2.

should then be expressly noted in the manuscript.[83] Preregistration (see above) can help with this process because it requires that researchers think through data collection before it occurs. For researchers collecting personal information, privacy will often be the most pertinent limitation.[84] Privacy issues can in many cases be managed through seeking consent to release data through the consent procedure and through de-identifying data, both of which should be approved the relevant institutional review board. The latter should be undertaken carefully and in accordance with applicable rules. In some cases, de-identification may not be possible to the extent necessary for ethical sharing (e.g., when risks of re-identification are high). Fortunately, repositories exist where access to raw data can be protected by qualified personnel,[85] and best practices exist for sharing data from human research participants.[86]

Analytic code, such as the scripts produced by several statistical software packages, allow readers to understand how the data produced the reported findings. Statistical packages typically allow the author to annotate the code with plain language explanations.[87] Sharing code makes it possible for others to verify conclusions. Indeed, users of the research should not simply be asked to trust that the reported conclusions are accurate; they should have the opportunity to verify those conclusions for themselves.[88]

---

[83] Richard D. Morey et al., *Peer Reviewers' Openness Initiative: incentivizing open research practices through peer review*, 3 R. SOC. OPEN SCI. 150547 (2016) at 5-6.

[84] Klein et al., *A practical guide, supra* note 19 at 4.

[85] Center for Open Science, *Approved Protected Access Repositories*, https://osf.io/tvyxz/wiki/8.%20Approved%20Protected%20Access%20Repositories/ (accessed 2020).

[86] See Klein et al., *A practical guide (supplementary material)*, *supra* note 81 at "Anonymization". See generally Michelle N. Meyer, *Practical Tips for Ethical Data Sharing*, 1(1) ADV. METH. & PRACT. PSYCHOL. SCI. 131 (2018).

[87] R Studio, *R Markdown*, https://rmarkdown.rstudio.com/ (accessed 2020); Jupyter, *Jupyter*, https://jupyter.org/ (accessed 2020).

[88] "The Royal Society's motto 'Nullius in verba' is taken to mean 'take nobody's word for it'. It is an expression of the determination of Fellows to withstand the domination of authority and to verify all statements by an appeal to facts determined by experiment." The Royal Society, *History of the Royal Society*, https://royalsociety.org/about-us/history/ (accessed 2020).

(c) Open your materials

Open materials also enhance a study's credibility. The OSF allows users to create a project page that contains data, analytic code, materials, manuscripts, and a brief wiki explaining the gist of the study. Researchers may wish to share materials like interview protocols and scripts, surveys, and any image, video, or audio files that were presented to participants.

One of the clearest credibility benefits of open materials is replicability.[89] In other words, future researchers can build off existing work, using materials (e.g., surveys) in different times and contexts. For example, a researcher may wish to repeat a survey or interview that was conducted in the past to see if there had been some change over time. And even for studies that are not quantitative, open materials contribute to a more efficient research ecosystem. For instance, if a researcher uploads interview questions to an open access repository, other researchers conducting a similar study can more easily build off that work, reusing questions that they think would be useful in their project.

(d) Consult an existing or analogous checklist when possible

We are not aware of any law journals that require or encourage authors to complete checklists when submitting their work for publication. There may, however, be an existing checklist for any given methodology being used by an empirical legal researcher, developed by others using the same methodology. For instance, a group of social and behavioral scientists recently created a checklist for conducting transparent research in their field.[90] They used an iterative, consensus-based protocol (i.e., a "reactive-Delphi" process) to help ensure that the checklist reflected the views of a diverse array of researchers and stakeholders in their field. Empirical legal researchers conducting behavioral research may find this transparency checklist

---

[89] Nosek & Errington, *supra* note 11.
[90] Aczel et al., *supra* note 51.

useful when planning their research and preparing it for publication. The Equator Network also curates a database of reporting checklists relevant to various research methods and disciplines.[91]

2. Three specific applications of research credibility to ELR

(a) Case law analysis

Empirical case law analysis has been widely used to address important legal issues.[92] As with other methods, credible research practices can be used to strengthen the inferences drawn from case law analysis and help ensure their enduring usefulness and impact.[93] In this subsection, we will drill down into two specific ways credible practices can be applied to empirical analyses of legal authorities: preregistering these studies, and using transparent methods to conduct them.

Preregistration poses a particular challenge when, as with analysis of legal authorities, the data are pre-existing. This is because, in preregistration's purest form, the hypothesis and methodology should be developed before the researchers have seen the data.[94] If this is not the case, the researchers may inadvertently present their hypotheses as independent of the data, when they were inadvertently constructing an explanation for what they already (in part) knew. Another challenge is that researchers may be tempted to sample and code cases in a way that fits their narrative. For instance, researchers may unconsciously determine that cases are relevant or

---

[91] Equator Network, *supra* note 51; See also American Psychological Association, *Journal Article Reporting Standards (JARS)*, https://apastyle.apa.org/jars (accessed 2020).

[92] See JEFFREY A. SEGAL & HAROLD J. SPAETH, THE SUPREME COURT AND THE ATTITUDINAL MODEL REVISITED (Cambridge University Press 2002); Epstein & King, *supra* note 4; Nita A. Farahany, *Neuroscience and behavioral genetics in US criminal law: an empirical analysis*, 2(3) J. LAW BIOSCI. 485 (2015); Jennifer A. Chandler, *The use of neuroscientific evidence in Canadian criminal proceedings*, 2(3) J. LAW BIOSCI. 550 (2015); Paul Catley & Lisa Claydon, *The use of neuroscientific evidence in the courtroom by those accused of criminal offenses in England and Wales*, 2(3) 2(3) J. LAW BIOSCI. 510 (2015); C. H. de Kogel & E. J. M. C. Westgeest, *Neuroscientific and behavioral genetic information in criminal cases in the Netherlands*, 2(3) J. LAW BIOSCI. 580 (2015).

[93] Epstein and King, *supra* note 4 at 38-45, urged researchers to use more reproducible practices for this reason. We agree and will try to make this general exhortation more concrete. See generally Chalmers & Glasziou, *supra* note 10

[94] Nosek et al., *The Preregistration Revolution*, *supra* note 34 at 3-4 under "Challenge 3: Data Are Preexisting".

irrelevant for their sample based on what would produce a more publishable result. While these challenges are significant, they do not mean that preregistration is not worthwhile in case law analysis. Indeed, other useful methods like systematic reviews and meta-analyses use pre-existing data, but also incorporate preregistration as part of best practices.[95]

One of us has some experience preregistering case law analyses and has found it to be a challenging but useful exercise.[96] In a recent study, for instance, he and colleagues sought to determine whether a widely-celebrated Supreme Court of Canada evidence law case had, in fact, produced more exclusions of expert witnesses accused of bias than the previous doctrine had allowed for.[97] The challenge was that it would have been useful to take a look at some cases before coding them to understand how long the process would take (e.g., for staffing purposes) and how to set up the coding scheme (e.g., would judges clearly advert to different aspects of the new doctrine so that the coders could unambiguously say courts were relying on these rules?). However, they were also aware that it would be easy to change their coding scheme and inclusion criteria based on the data to show a more startling result (e.g., shifting the timeframe slightly might make it seem as if the focal case had more or less of an impact). In other words, some type of preregistration was needed, but the standard form seemed too restrictive.

Accordingly, they decided to establish the temporal scope of the search prior to looking at the cases, but to accommodate the difficulty of pre-deciding on the criteria by reading a portion of the cases prior to establishing the coding scheme. They disclosed this encroachment into the data in the preregistration, so that readers could adjust their interpretation of the results

---

[95] PRISMA Transparent Reviews of Systematic Reviews and Meta-Analyses, *Registration*, http://www.prisma-statement.org/Protocols/Registration (accessed 2020).

[96] Jason M. Chin, Michael Lutsky & Itiel E. Dror, *The Biases of Experts: An Empirical Analysis of Expert Witness Challenges*, 42(4) MAN. L. J. (2019) at 35, with the preregistration at https://osf.io/n4j68/registrations (accessed 2020); see also Jason M. Chin et al., *IMM v The Queen Retrospective*, https://osf.io/vam7t/registrations (accessed 2020).

[97] Chin et al., *supra* note 96.

accordingly.[98] This was useful in that they were able to account for several issues that would

have been difficult to anticipate without some prior knowledge of the cases. For instance,

sometimes judges in bench trials would not exclude a witness, but rather say that the witness

would be assigned no weight. It was hard to determine if this should be coded as an exclusion.

By looking at some of the data, they were able to anticipate this for the bulk of the cases. Had it

been done completely ad hoc, without any preregistration, it would have been difficult to make

the decision about how to code these cases in an unbiased way. They also took the step of

disclosing cases that were borderline and required discussion amongst the authors, but did not do

so as systematically as they would in the future.[99] In the end, they were able to give what they

thought was a credible picture of the target case's effect, with the preregistration helping to

reduce the possible influence of bias and helping to highlight the study's limitations.

Other transparency and openness efforts can also improve case law analysis. One method

now common in systematic reviews (which also use pre-existing data) that legal researchers may

leverage are "PRISMA diagrams" (see Figure 3).[100] These diagrams thoroughly document the

process by which an existing body of knowledge is searched. Like case law researchers,

systematic reviewers typically start with keyword searches to identify a group of published

studies. These are then winnowed down, based on preregistered criteria, to what is ultimately

analyzed. The resulting PRISMA diagram provides a concise explanation of that process.

To see the value of a PRISMA diagram in the context of ELR, consider a recent series of

studies that looked at the use of neuroscience in courts across several jurisdictions, finding

---

[98] Chin et al., *The Biases of Experts*, *supra* note 96, *preregistration*.
[99] See *id.* at fn. 85.
[100] PRISMA, *PRISMA 2009 Flow Diagram*, http://prisma-statement.org/documents/PRISMA%202009%20flow%20diagram.pdf (accessed 2020).

reliance on neuroscience evidence has increased and detailing its different uses.[101] Subsequent researchers may wish to extend those findings to see whether the discovered trends and uses have changed. They may also want to stand on the shoulders of the earlier researchers and extend the analysis to other jurisdictions. In either case, they would need clear methods to follow in order to reproduce the searches, exclusion criteria, and coding. However, we have noticed considerable heterogeneity in the way methodologies were described in those studies, and in similar ones.[102] Following a well-understood framework like PRISMA to see what exactly was searched and how the search list was reduced down to what was presented in the article would be beneficial.

(b) Survey studies

Legal researchers have used surveys to address a host of questions, like lawyers' reports on their well-being,[103] judges' attitudes towards evidence procedures,[104] and the public's views on what is a violation of privacy.[105] In fact, almost half of all quantitative studies on topics related to crime and criminal justice use surveys.[106]

There are many existing resources to help improve survey methodology generally (e.g., sampling, wording of questions and prompts).[107] We, however, are interested in improving the

---

[101] Farahany, *supra* note 92; Chandler, *supra* note 92; Catley & Claydon, *supra* note 92; de Kogel & Westgeest, *supra* note 92.

[102] For example, one listed keywords searched, without explaining if the list was exhaustive and how variations of those words were handled: Farahany, *supra* note 92 at 490. Another provided precise search terms and the reasons groups of cases were excluded, Chandler, *supra* note 92 at 553-5. Both methods are preferable to those that would just provide examples of keywords used, see Gary Edmond, *Latent Science: A History of Challenges to Fingerprint Evidence in Australia* 38(2) U. Q. L. J. 301 (2019) at fn. 4: "Databases were searched using terms including...".

[103] John Monahan & Jeffrey Swanson, *Lawyers at Mid-Career: A 20-Year Longitudinal Study of Job and Life Satisfaction*, 6(3) J. EMPIR. LEG. STUD. 451 (2009).

[104] Stephanie Domitrovich, Mara L. Merlino & James T. Richardson et al., *State trial judge use of court appointed experts: Survey results and comparisons*, 5(3) JURIMETRICS 371 (2010).

[105] Matthew B. Kugler & Thomas H. Rousse, *The Privacy Hierarchy: Trade Secret and Fourth Amendment Expectations*, 104 IOWA L. REV. 1223 (2019) at 1249-63.

[106] Gary Kleck, Jongyeon Tark & Jon J. Bellows, *What methods are most frequently used in research in criminology and criminal justice?*, 34 J. CRIM. JUST. 147 (2006).

[107] See e.g., PETER M. NARDI, DOING SURVEY RESEARCH (Routledge 2015).

credibility of survey research in law – conducting studies such that their findings are reproducible, errors are detected and corrected, and conclusions are calibrated to the strength of the evidence.

One key aspect of credible survey research – and one that is regularly breached in law and elsewhere – is data transparency. In criminology, for instance, closed data practices protracted one incident over years, in which co-authors of a series of studies repeatedly tried and failed to obtain raw data from the lead author, as well as evidence that the reported surveys were in fact conducted.[108] In the case of surveys, data transparency pertains to reporting and making available to other researchers all key information about the questionnaire and data collection procedures. An excellent guide is the American Association of Public Opinion Research's (AAPOR) survey disclosure checklist, which recommends researchers disclose the survey sponsor, data collection mode, sampling frame, field date (or dates of administration), and exact question wording.[109]

Beyond the exact question wording, which should be disclosed per AAPOR, we also recommend researchers make the entire questionnaire itself publicly available. This permits others to not only reuse the questions, but also to replicate the question ordering, which can have large effects on responses.[110]

---

[108] Justin T. Pickett, *The Stewart Retractions: A Quantitative and Qualitative Analysis*, 17(1) ECON. J. WATCH 152 (2020).

[109] American Association of Public Opinion Research, *Survey Disclosure Checklist*, https://www.aapor.org/Standards-Ethics/AAPOR-Code-of-Ethics/Survey-Disclosure-Checklist.aspx (accessed 2020). See also recent openness and transparency guides for survey research and mail data collection: Ines Schaurer, Tanja Kunz, and Tobias Heycke, *Documentation of online surveys*, https://www.gesis.org/fileadmin/upload/SDMwiki/2020_schaurer_online_1.pdf (accessed 2020); Sven Stadtmüller and Christoph Beuthner, Documentation of mail data collection, https://www.gesis.org/fileadmin/upload/SDMwiki/2020_stadtmueller_maildocumentation_1.pdf (accessed 2020).

[110] DON A. DILLMAN, JOLENE D. SMYTH & LEAH MELANI CHRISTIAN, INTERNET, PHONE, MAIL, AND MIXED-MODEL SURVEYS: THE TAILORED DESIGN METHOD (4TH ED 2014).

In terms of sampling, researchers should state explicitly whether sample selection was probabilistic (i.e., using random selection), in addition to describing how sampled respondents compare to the population of interest. This is important because researchers are increasingly using online non-probability samples, recruited via crowdsourcing or opt-in panels,[111] but are mislabeling these samples as "representative" when they match the general population on a handful of chosen variables (e.g., gender, race). Mischaracterizing these online convenience samples as probabilistic samples may lead readers to put more confidence in the generalizability of findings than is warranted. It also obscures the fact that non-random sampling necessitates different types of inference. That is, even when non-probability samples look similar to the population on a few chosen demographic variables, inferences from them still depend on assumptions and statistical adjustments (i.e., model-based inference), rather than probability theory (i.e., design-based inference).

Another key type of information that researchers using surveys should disclose is the response rate. Reviews of the literature have revealed widespread failure to report response rates and inconsistencies in calculating those rates.[112] Smith concluded that "disclosure standards are routinely ignored and technical standards regarding definitions and calculations have not been widely adopted in practice."[113] The problem has only worsened with the expansion of online sampling, which has further complicated the calculation of responses. For example, correctly calculating the cumulative response rate (CUMR) for a survey fielded with a probability-based online panel, like NORC's AmeriSpeak panel, requires taking into consideration the initial panel

---

[111] Andrew Thompson & Justin T. Pickett. *Are relational inferences from crowdsourced and opt-in samples generalizable? Comparing criminal justice attitudes in the GSS and five online samples*, J. QUANT. CRIMINOL. (2019), https://link.springer.com/article/10.1007/s10940-019-09436-7 (accessed 2020).
[112] Tom W. Smith, *Developing nonresponse standards*, IN SURVEY NONRESPONSE 27-41 (Robert M. Groves et al. eds., 2002).
[113] *Id*. at 39.

recruitment rate (RECR) and the panel profiling rate (PROR), yet researchers frequently misreport the study-specific completion rate (COMR) as the response rate.[114] A path forward is to require all researchers using survey data to report the response rate and to adhere to the AAPOR's Standard Definitions when calculating that rate.[115]

Finally, survey researchers should transparently document and disclose their selection criteria. Beyond common eligibility criteria, such as adult status and country of residence, online platforms give researchers numerous other options, which can impact sample composition, and which are rarely reported. On Amazon *Mechanical Turk*, for example, researchers commonly restrict participation to workers with certain reputation scores (e.g., at least 95% approval) and/or prior Human-Intelligence-Task (HIT) experience (e.g., must have completed 500 prior HITs).[116] Such eligibility restrictions can have a profound effect on data quality and sample composition.[117] Therefore, we recommend that researchers using online platforms, like Amazon's Mechanical Turk, disclose all employed eligibility restrictions. Additionally, researchers should disclose if they exclude respondents for quality control reasons, such as for speeding through the survey, failing attention checks, or participating repeatedly (e.g., duplicate Internet Protocol addresses), and they should report how the exclusions affect findings. All such exclusions, if undisclosed and decided on after looking at effects on findings, would amount to questionable research practices and inflate the false positive rate.

---

[114] The correct calculation is: CUMR = RECR x PROR x COMR.

[115] American Association of Public Opinion Research, *Standard Definitions*, https://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf (accessed 2020).

[116] KIM BARTEL SHEEHAN & MATHEW PITTMAN, AMAZON'S MECHANICAL TURK FOR ACADEMICS: THE HIT HANDBOOK FOR SOCIAL SCIENCE RESEARCH, (Melvin & Leigh 2016).

[117] Eyal Peer, Joachim Vosgerau & Alessandro Acquisti, *Reputation as a sufficient condition for data quality on Amazon Mechanical Turk*, 46 BEHAV. RES. METHODS 1023 (2014).

(c) Qualitative research

Qualitative methods also play an important role in ELR.[118] These include methods like ethnographies,[119] interviews and focus groups,[120] and case studies.[121] While it may be tempting to think that the reforms we have discussed are inappropriate for seemingly more freeform and exploratory research, we suggest that is not the case. In saying this, we do not deny that there are fundamental differences between quantitative and qualitative methods. For instance, Lisa Webley has noted that qualitative researchers often see their work as more interpretivist than positivist, more inductive than deductive, and, at times, more interested in socially constructed facts than those that purport to have universal meaning.[122] Additionally, many qualitative researchers are skeptical of the modern open science movement, which they see as imposing quantitatively-focused evaluation criteria on qualitative researchers without understanding the contextual or epistemological differences between these types of research.[123] None of these epistemological differences, however, undermine the importance of research credibility. Rather, as we will discuss, qualitative researchers can leverage many existing reforms, even though many are grounded in positivist frameworks, to make their work easier to access and to help ensure its long-term impact.

---

[118] Lisa Webley, *Qualitative Approaches to Empirical Legal Research*, IN THE OXFORD HANDBOOK OF EMPIRICAL LEGAL RESEARCH 927-49 (Peter Cane & Herbert M. Kritzer eds., 2012).

[119] Sally Falk Moore, *Law and Social Change: The Semi-Autonomous Social Field as an Appropriate Subject of Study*, 7 LAW & SOC'Y REV. 719 (1973).

[120] Zoe Rathus, *The Research Says: Perceptions on the Use of Social Science Research in the Family Law System*, 46 FED. L. REV. 85 (2018).

[121] Allison Christians, *Critical Issues in Comparative & International Taxation: Case Study Research and International Tax Theory*, 55 St. Louis Univ. L. J. 311 (2010).

[122] Webley, *supra* note 118 at 929-31.

[123] Katherine A. Tamminen & Zoë A. Poucher, *Open science in sport and exercise psychology: Review of current approaches and considerations for qualitative inquiry*, 36 PSYCHOL. SPORT EXERC. 17 (2018).

Before delving into the reforms, we note that we are not the first to highlight the importance of credibility in qualitative legal research.[124] For instance, Allison Christians examined research methodology in a meta-analysis of case studies in international law. She found that not one article explained why the specific case was chosen: "In each case, the articles simply identify the event or phenomenon as a 'case' without further discussion."[125] This, as she goes on to note, raises the possibility of selection bias, whereby the case is not representative or probative of the claim it seeks to support. Further, Christians found that studies did not sufficiently explain why certain material was collected to document the case (and why other material was left out): "What is missing from the literature and what might make the data even more compelling, is a discussion about the authors' objectives, processes, and reasoning for collecting and using the data…"[126] And, when data and materials were relied on, Christians found that these sources were often not cited.[127] Recall that data citation is a TOP guideline that only 2 law journals in our sample addressed (and very weakly so, see Table 1).

Christians' observation about thinking through and justifying case selection reinforces the importance of preregistration of qualitative studies (when preregistration aligns with the research epistemology).[128] Indeed, preregistration is an increasingly discussed reform in qualitative research. One group recently completed a preregistration form through a consensus-based process (i.e., the same process used to create the checklist for behavioral studies above) that

---

[124] Christians, *supra* note 121; Webley, *supra* note 118 at 935: "Also important are the extent to which she is willing to pilot her method, to make adjustments in the light of the pilot, to be reflexive and to report on the strengths and weaknesses of her research".

[125] Christians, *supra* note 121 at 336-7.

[126] *Id*. at 362. Similarly, she notes: "None of the international tax case studies includes a description of the author's reasoning regarding how the case is or should be constructed.": *Id*. at 356.

[127] *Id*. at 359: "These authors—perhaps like many legal scholars—used their discussions with these individuals to better understand the studied subject or to construct theories about the studied subject, but they did not cite to the primary source of data—namely, notes from interviews or e-mail correspondence".

[128] Tamarinde L. Haven & Leonie Van Grootel, *Preregistering qualitative research*, 26(3) ACCOUNT. RES. 229 (2019).

many researchers in the field participated in.[129] Qualitative preregistrations may include details about the research team's background as it relates to the study (as a form of reflexivity, i.e., attendance to the experiences, positions, and potential biases the researchers bring to bear on what they are studying), research questions and aims, planned procedures, sampling plans, data collection procedures, planned evidence criteria, and triangulation, auditing, and validation plans. As qualitative research tends to be more iterative than quantitative research, preregistrations may be most useful not as a means for researchers to try to establish "objectivity," but rather as a means for researchers to fully explore assumptions they may be making going into their study, and another tool for reflexivity as the study progresses.

If preregistration does not align with one's research epistemology, it is still possible to engage in transparent practices so others may evaluate research decisions and learn from researchers' practices. Researchers may be interested in maintaining open laboratory notebooks (adapted to qualitative practices)[130] and/or sharing their research materials (e.g., recruitment materials, interview and focus group protocols, fieldnotes, codebooks, etc.) on a repository like the OSF. Data may also be shared on the OSF or, for instance, the Qualitative Data Repository.[131] There are, however, important ethical considerations to account for before sharing data. Kristen Monroe outlined several concerns with the Data Access and Research Transparency (DA-RT) and Journal Editors Transparency Statement (JETS) initiatives as they relate to qualitative research.[132] These include: space constraints that may hinder full accounting for qualitative data, participant protection, the time needed to prepare qualitative data for sharing,

---

[129] Tamarinde Haven et al., *Preregistering Qualitative Research: A Delphi Study*, https://osf.io/preprints/socarxiv/pz9jr/ (accessed 2020).

[130] Matthieu Schapira, Open Lab Notebook Consortium & Rachel J. Harding, *Open laboratory notebooks: good for science, good for society, good for scientists*, 8(87) F1000RESEARCH (2020).

[131] Qualitative Data Repository, *The Qualitative Data Repository*, https://qdr.syr.edu/ (accessed 2020).

[132] Kristen Renwick Monroe, *The Rush to Transparency: DA-RT and the Potential Dangers for Qualitative Research*, 16(1) PERSPECT. POLITICS 141 (2018).

costs of data collection, the right of first usage, and a potential chilling effect on certain research topics. Others have outlined concerns surrounding missing layers of interpretation and the importance of consent as an ongoing process.[133]

Researchers should handle datasets involving information from vulnerable populations, for example sexual assault survivors or refugees, with care, such that participants' personal information is appropriately protected. Fortunately, many data repositories do offer access controls such that researchers may embargo data or provide conditions for access, if desired. Additionally, some researchers have begun to include consent language around sharing data with other researchers on the condition the participants' anonymity is preserved or offering conditional consent, where participants can agree to participate but not share data with anyone other than the study authors. It is also important to note the same documents that make up "audit trails" (e.g. field notes, interview and focus group protocols, etc.) are useful tools for making qualitative research more open and transparent, and may be particularly beneficial for those learning how to conduct qualitative studies.[134]

Part IV. Guidance for journals and institutions that wish to encourage credible research practices

Most researchers readily endorse norms associated with the reforms we have discussed above.[135] Still, in other fields, expressed acceptance of norms exceeds the actual behaviors they are associated with (e.g., researchers say sharing materials is important, but do not always live up that ideal).[136] In Part III, we attempted to address one reason behaviors may be lagging behind

---

[133] Peter Branney et al., *A context-consent meta-framework for designing open (qualitative) data studies*, 16(3) QUAL. RES. 483 (2019) https://doi.org/10.1080/14780887.2019.1605477.

[134] Alexander C. Tsai et al., *Promises and pitfalls of data sharing in qualitative research*, 169 SOC. SCI. MED. 191 (2016).

[135] Anderson, Martinson & De Vries, *supra* note 65.

[136] *Id*.

norms – a general lack of concrete guidance aimed at legal researchers. We will now address two more factors that affect the behavior of researchers: incentives and policies. We will draw on general research on these factors, but adapt them to the distinctive ecosystem of legal research and teaching.

1. Journals

As we saw above, there is considerable heterogeneity in journal guidelines among law journals, both in the student-edited journals and the small number of peer-reviewed journals we considered. The most significant advancement in transparency that we found was in two journals adopting guidelines from economics journals. To us, this suggests that journal editors and boards in the ELR space may be open to adopting new guidelines, but that it is important that the task be as easy as possible, and the guidelines be tested in similar fields. For that reason, we will suggest low-cost and pre-vetted moves journals may make.

Peer-reviewed journals should consider the sample TOP implementation language curated by the COS.[137] These can be adapted to the needs of the specific legal research journal. Similarly, journals may consult the guidelines of other ELR journals in our sample that have implemented TOP (Table 1). Registered Reports – which have been adopted by journals in fields ranging from psychology to medicine – can be fairly easily rolled out by law journals, with recommended author and reviewer instructions available for re-use.[138] Registered reports may be especially attractive in law where early-stage review may assist researchers with more limited backgrounds in methodology. Looking towards the horizon, it may be time for the field to

---

[137] Center for Open Science, TOPMixedLevelsJournals.gdoc, https://osf.io/edtxm/ (accessed 2020); For guidelines more relevant to qualitative research, see Tsai et al., *supra* note 134.

[138] Center for Open Science, *Registered Reports: Peer review before results are known to align scientific values and practices*, under "Resources for Editors", https://www.cos.io/our-services/registered-reports (accessed 2020).

develop a journal with a philosophy that values methods over results and that regularly publishes work about methodology itself (from psychology, potential models are *Advances in Methods and Practices in Psychological Science* and *Collabra: Psychology*).

The situation with student-edited journals is more complicated because, among other reasons, their editorial boards experience a great deal of turnover, they may be less likely to have empirical backgrounds, and, as Part III indicates, current guidelines have the most room for improvement. The landscape in student-edited journals also seems to be more competitive, with editors taking into account acceptance of the article at other eminent journals. The culture of concurrently submitting to many journals places time pressure on student editors. As a result, they may be hesitant to screen articles that, despite seeming impressive in some ways, do not meet high methodological standards.

These unique hurdles at student law reviews are not insurmountable in the long run. One incremental measure may be for these journals to award badges. Recall that badges do not necessarily factor into article acceptance, but instead allow authors to signal to others that they have taken steps to improve their work's credibility and usability.[139] More generally, note that student-edited journals are not immune from change. About 15 years ago, many signed on to an agreement to accept articles with fewer words.[140] Some of the authors of the current article are beginning a project to draft sample guidelines designed for law journals that occasionally publish empirical work. Having these ready-made guidelines may make the change less daunting. We

---

[139] Kidwell et al., *supra* note 49. Center for Open Science, *2. Awarding Badges*, https://osf.io/tvyxz/wiki/2.%20Awarding%20Badges/ (accessed 2020).

[140] Michael C. Dorff, *Thanks to a Joint Statement by Top Law Journals, Law Review Articles Will Get Shorter, But Will They Get Better?*, https://supreme.findlaw.com/legal-commentary/thanks-to-a-joint-statement-by-top-law-journals-law-review-articles-will-get-shorter-but-will-they-get-better.html (accessed 2020).

plan to circulate them to student-edited journals along with many of the justifications presented in this article. Please contact the corresponding author if you would like to contribute.

Finally, the simple step of encouraging the submission of replication studies can be an important step toward improvement and reform. Without empirical evidence about a field's replicability, it may be challenging to see the need for reform. Either individual studies or larger efforts meant to more systematically estimate the replicability of a sub-discipline can provide insight into the extent and consequences of these problems.

2. Law schools and faculties

Law schools and faculties can also play a role in encouraging credible practices. This naturally begins with hiring, where committees already seem to place some value on empirical research by hiring individuals who do such work.[141] It is less clear, however, whether these committees place much value on the credibility and rigor of empirical work (as opposed to factors like its surprisingness and ability to draw headlines). If committees do not take credible practices into account, then hiring practices may perpetuate irreproducible research.

Hiring committees may wish to align their search criteria and candidate evaluation with recent work laying out frameworks for basing researcher assessment on the credibility of their work.[142] The Hong Kong Principles distill research assessment into five factors. They seek to move fields from success indicators like the esteem of journals and impact factors to other criteria, like the dissemination of knowledge through open data and the analysis of existing, but poorly understood work, through evidence synthesis and meta-analysis.[143]

---

[141] Diamond, *Empirical Legal Scholarship: Observations on Moving Forward*, *supra* note 61 at 1229.
[142] David Moher et al., *Assessing scientists for hiring, promotion, and tenure*, 16(3) PLoS BIOL e2004089 (2018); David Moher et al., *The Hong Kong Principles for assessing researchers: Fostering research integrity*, 18(7) PLoS BIOL e3000737 (2020).
[143] Moher et al., *Hong Kong Principles*, *Id*.

Precedents are available to assist institutions seeking to change their hiring practices. The COS maintains a list of job listings that refer to research practices.[144] For instance, a recent University of Toronto listing stated: "Our department embraces the values of open science and strives for replicable and reproducible research. We therefore support transparent research with open data, open material, and pre-registrations. Candidates are asked to describe in what way they have already pursued and/or plan to pursue open science."[145] These principles may also be applied to tenure standards.

After hiring, more may be done, as some have suggested,[146] to promote collaboration between those who have specialized empirical training and experience, and those who do not.[147] One barrier to this initiative is authorship norms, and the concern that the methodological work may go unrecognized. In these circumstances, law schools may take note of a move towards a contributorship model of authorship, which recognizes the various types of work that go into a publication.[148]

Internal encouragement within schools and faculties can only go so far, especially when the broader environment rewards high impact publication (in which impact is often not directly related to strength of methodology). This may especially be the case in the U.S., where law school rankings are so tied to the eminence of the journals that the faculty publishes in. Still, both in the U.S. and abroad, there are incentives to focus on methodology.[149] For example, in the US, one influential publication is considering using citation counts of individual researchers

---

[144] Center for Open Science, *Universities*, https://osf.io/kgnva/wiki/Universities/ (accessed 2020).
[145] *Id*.
[146] Richard L. Revesz, *A Defense of Empirical Legal Scholarship*, 69 U. CHI. L. REV. 169 (2002).
[147] See also Van Zandt, *supra* note 56.
[148] Alex Holcombe, *Farewell authors, hello contributors*, 571 NATURE 147 (2019).
[149] Willem E. Frankenhuis & Daniel Nettle, *Open Science Is Liberating and Can Foster Creativity*, 13(4) PERSP. PSYCHOL. SCI 439 (2018).

(rather than Journal Impact Factors) to assess the productivity of law schools.[150] With this in mind, legal researchers may be swayed by findings that sharing data is associated with increased citations.[151] Similarly, funders are increasingly concerned with and sometimes require open practices.[152] In other words, tradeoffs between the expectations of current ranking systems and research credibility may not be as stark as they seem at first blush.

Part V. Conclusion

In producing knowledge for the legal system, empirical legal researchers have a heightened duty to present the full picture of the evidence underlying their results. We are excited for what the next several years hold for better fulfilling that duty. While there are sticking points, like the need for more training and the distinctive situation with student-edited journals, there are also an increasing number of models to follow from cognate fields, and an energized group of researchers motivated to put them into action. In the past, calls for change in ELR have sometimes gone unheeded, but never before have they been made in the context of a large, sustained movement in the rest of the research ecosystem.

---

[150] Robert Morse, U.S News & World Report, *U.S. News Considers Evaluating Law School Scholarly Impact*, https://www.usnews.com/education/blogs/college-rankings-blog/articles/2019-02-13/us-news-considers-evaluating-law-school-scholarly-impact (accessed 2020).
[151] Heather A. Piwowar & Todd J. Vision, *Data reuse and the open data citation advantage*, 1 PEERJ e175 (2013).
[152] NASEM, *supra* note 7 at 129-30.

<u>Figure 1. Adoption of open and transparent practices by researchers in the social sciences</u>

Participants were asked the year they first engaged in one of the following practices (if they had): made their data available online, made their study instruments (i.e., materials) available online, and preregistered a study. Participants were researchers in psychology, economics, political science, and sociology.
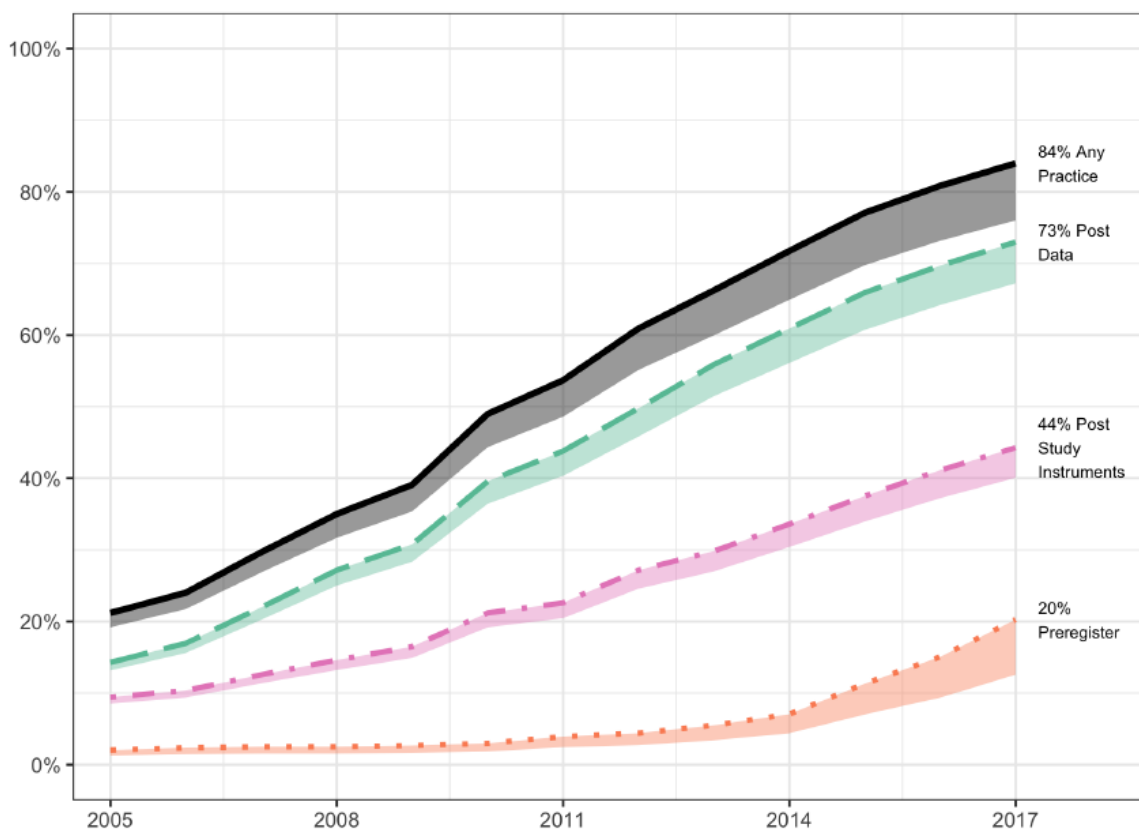
Figure 2. First listed hypothesis confirmed in standard and registered reports

The authors compared a sample of standard reports to a sample of registered reports. They found that 96.05% of standard reports found the first listed hypothesis was confirmed, whereas 43.66% of such hypotheses in registered reports were confirmed.
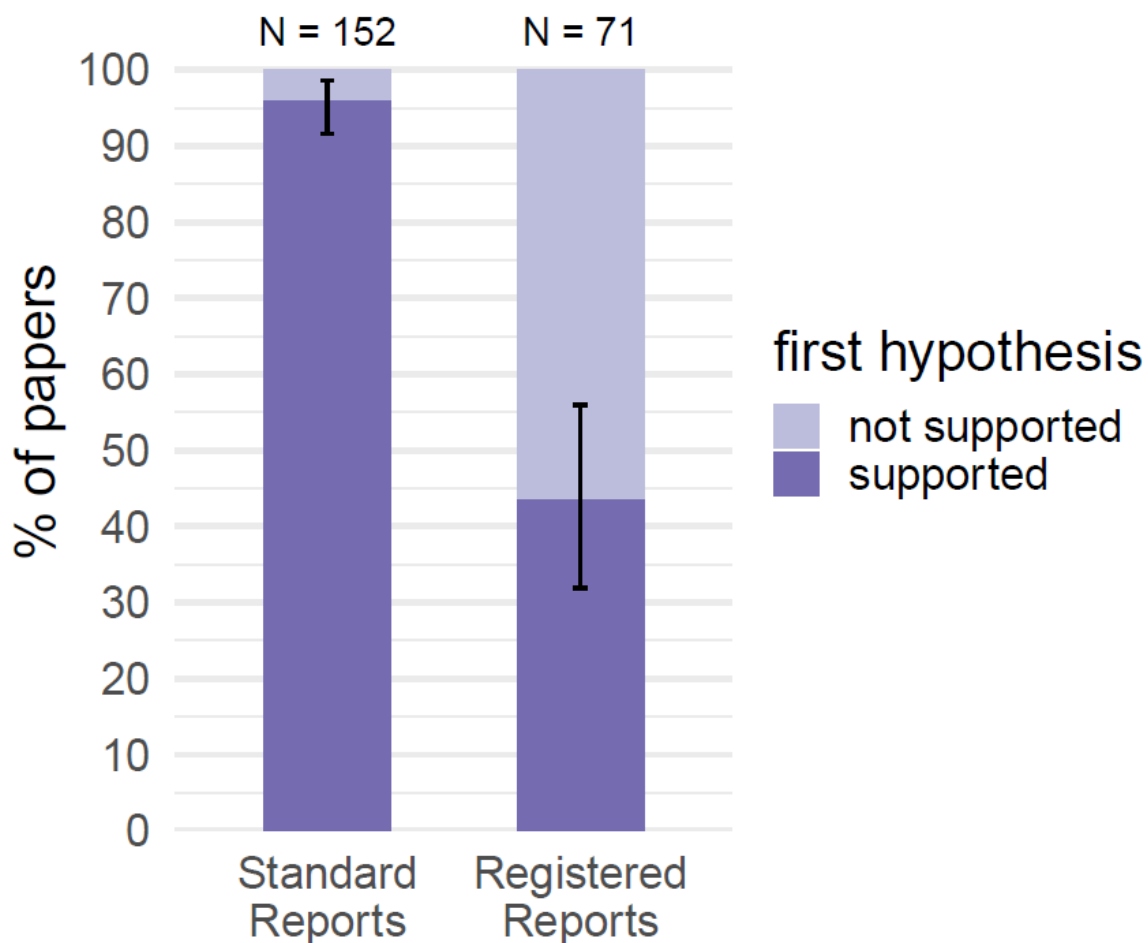
Figure 3. PRISMA Flow Diagram

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow

diagram is used in many fields to improve the transparency of the secondary analysis of pre-

existing studies. Specifically, it makes clearer why some studies were included in the analysis

and some were not. As demonstrated in this figure, it can be adapted by empirical legal

researchers to transparently report the cases included in the analysis and those that were

excluded. In this example, the researcher searched two prominent databases, but this can be

changed as needed.
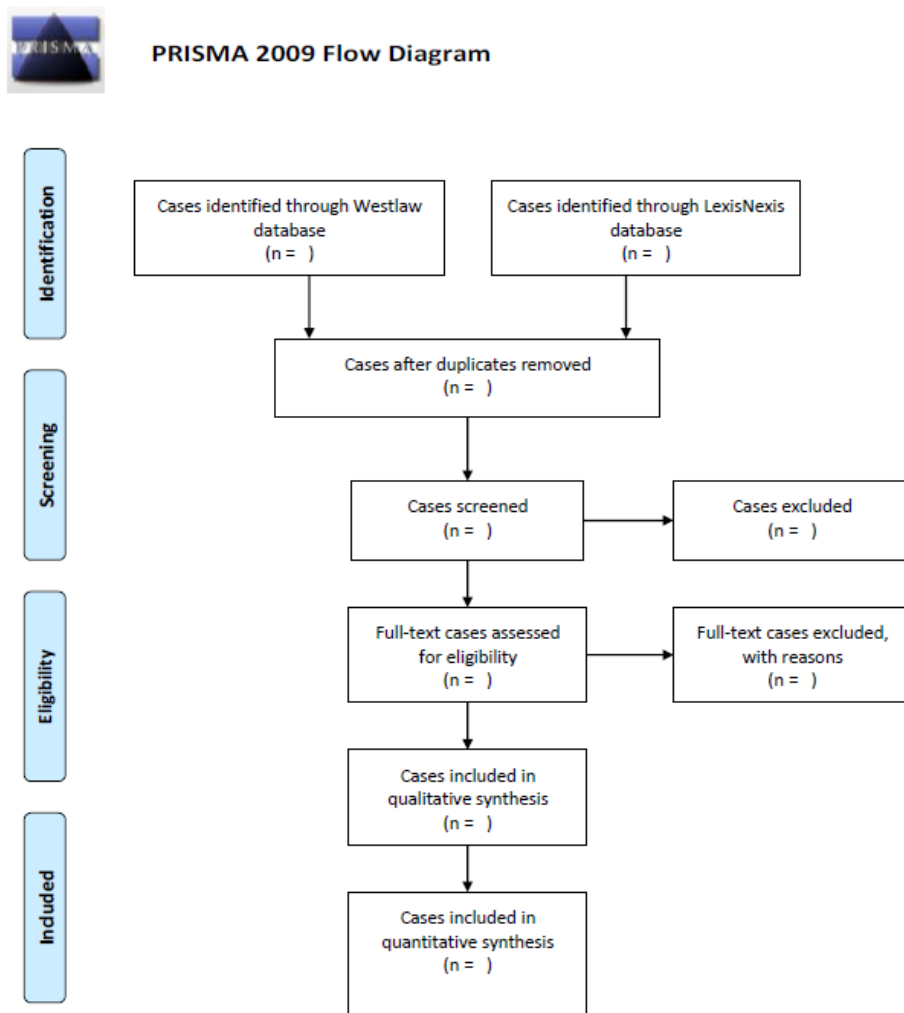
Table 1. Transparency and openness policies in law journals

Table 1. The top 25 journals according to the Washington & Lee list and 6 peer-reviewed empirical journals evaluated against the factors in the Transparency and Openness journal guidelines. TOP Factor is the sum of 10 items (https://osf.io/t2yu5/) that are awarded 0-3 based on how insistent the policy is: data citation, data transparency, analytic code transparency, materials transparency, reporting guidelines, study preregistration, analysis preregistration, replication, publication bias, and open science badges. The latter five items are not listed because all journals received 0s for them. The highest possible TOP Factor score is 30.

Journal	TOP Factor	Peer Reviewed?	Data Citation	Data Transparency	Analysis Code	Materials Transparency
Journal of Legal Studies	10	Yes	1	3	3	3
Journal of Law and Economics	10	Yes	1	3	3	3
Yale Law Journal	4	No	0	2	2	0
American Law and Economics Review	4	Yes	0	2	2	0
Stanford Law Review	2	No	0	2	0	0
Journal of Legal Analysis	2	Yes	2	0	0	0
New York University Law Review	2	No	0	2	0	0

California Law Review	0	No	0	0	0	0
Iowa Law Review	0	No	0	0	0	0
Michigan Law Review	0	No	0	0	0	0
Duke Law Journal	0	No	0	0	0	0
Georgetown Law Journal	0	No	0	0	0	0
University of Pennsylvania Law Review	0	No	0	0	0	0
William and Mary Law Review	0	No	0	0	0	0
Boston University Law Review	0	No	0	0	0	0
Texas Law Review	0	No	0	0	0	0
Minnesota Law Review	0	No	0	0	0	0
Harvard Law Review	0	No	0	0	0	0
Southern California Law Review	0	No	0	0	0	0

Journal of Empirical Legal Studies	0	Yes	0	0	0	0
Boston College Law Review	0	No	0	0	0	0
UCLA Law Review	0	No	0	0	0	0
University of Chicago Law Review	0	No	0	0	0	0
Supreme Court Review	0	No	0	0	0	0
Cornell Law Review	0	No	0	0	0	0
Columbia Law Review	0	No	0	0	0	0
Fordham Law Review	0	No	0	0	0	0
Law Probability and Risk	0	No	0	0	0	0
Notre Dame Law Review	0	No	0	0	0	0
Virginia Law Review	0	No	0	0	0	0

Table 2. A selection of data repositories for empirical legal research

Table 2. A list of six open access data repositories that empirical legal researchers may consider using. All but one (figshare) are non-profits. As described below, they provide hosting in a variety of countries and jurisdictions.

Name	Website	Highlights
figshare	https://figshare.com/	Free up to 100GB. Often used for sharing data analysis and code. Hosted in the UK. It is a for-profit company.
GESIS datorium	https://data.gesis.org/sharing/	Free. Hosted by the Leibniz Institute for the Social Sciences (Germany), a non-profit organization.
Open Science Framework	https://osf.io	Free. Also provides forms for preregistration that can be connected to projects. Operated by the Center for Open Science with options to store data in multiple countries. The Center for Open Science is a non-profit in the U.S.
The Qualitative Data Repository	https://qdr.syr.edu	Pay (waivers available). Operated by Syracuse University. Focuses on qualitative data and offers assistance in determining what to share, and how.
UK Data Service standard archiving	https://www.ukdataservice.ac.uk/	Free. Focuses on large-scale survey data. Data must be sent to the service, which also curates the data (i.e., self-deposit is not possible). The UK Data Service is funded by the Economic and Social Research Council, which is a public body.
Zenodo	https://zenodo.org/	Free. Operated by European Organization for Nuclear Research (CERN), a non-profit European research organization, and is hosted in the EU.