

Increasing the credibility of political science research: A proposal for journal reforms

Brendan Nyhan
Dept. of Government
Dartmouth College
nyhan@dartmouth.edu

September 11, 2014

Abstract

In this white paper, I propose a series of reforms intended to improve the rigor, timeliness, and replicability of published political science research. These changes could be proposed as part of the next solicitation for editors of the *American Political Science Review* or other disciplinary journals. If implemented, they would make our discipline a leader in publishing practices intended to minimize publication bias, encourage rapid and high-quality reviews, and maximize replicability and fairness to authors. In this way, they could help to enhance the quality and pace of the profession's scholarly output and improve perceptions of its rigor both within and outside of academia over time, increasing trust in political science findings and building the credibility that is necessary for successful public engagement.

This white paper was written for the American Political Science Association Task Force on Public Engagement. I thank Deborah Brooks, John Carey, Brian Greenhill, Jeremy Horowitz, Arthur Lupia, and Ben Valentino for helpful comments and discussion. The contents of this article are based in part on Nyhan (2012a,b).

How can political science improve its standing in and contribution to public life? While our discipline should engage more effectively in the national debate (see the Nyhan, Sides, and Tucker white paper in this report), we can also strengthen the quality and timeliness of the scholarship that we hope to bring to a wider audience.

In this white paper, I propose two sets of reforms to American Political Science Association (APSA) journal practices intended to achieve those goals.¹ The first proposal seeks to counter publication bias and increase incentives for careful research practices that generate replicable results, while the second is designed to improve the timeliness and quality of reviewer and editorial decisions during the review process.

If implemented carefully, these reforms could make the publication process more rigorous and efficient and thereby enhance the value of the findings we produce, which would be more timely and credible and therefore more valuable to the external audiences we hope to engage. In this way, improving our journal practices could enhance external perceptions of the field's value and the level of interest in and attention to our scholarship.²

Countering publication bias in political science

Academics face intense pressure to publish statistically significant findings in top journals as authors and to reject articles that fail to find such results as reviewers and editors. In practice, those incentives create extensive publication bias in disciplines including political science (Gerber and Malhotra 2008a; Gerber et al. 2010; Esarey and Wu N.d.), sociology (Gerber and Malhotra 2008b), economics (Doucouliagos 2005), and psychology (Ferguson and Heene 2012; Masicampo and Lalande 2012), resulting in published literatures that contain far too many narrowly statistically significant findings to be attributable to chance (i.e., heaping of p -values at just under $p < .05$). All too often, such findings are fragile to minor perturbations such as variations in model specification (Montgomery and Nyhan 2010) and fail to replicate in future studies (e.g., Donnellan, Lucas, and Cesario N.d.; Ritchie, Wiseman, and French 2012). By contrast, null findings are typically relegated to the so-called file drawer and excluded from the published record. Social scientists tend to think of medical and scientific journals as being more rigorous, but those fields appear to suffer from the same problems (Ioannidis 2005).

While some fraud may occur, the problem is more likely to be one of self-deception — as authors, reviewers, and editors, we are simply too good at rationalizing choices that produce the results we want, including p -hacking (Simmons, Nelson, and Simonsohn 2011; Simonsohn, Nelson, and Simmons 2013), the use of low-power studies (Button et al. 2013), and more. Moreover, our acceptance of these norms often leads to reviewer insistence on statistical significance or declining to publish articles that fail to reach this threshold based on other rationales (e.g., *post hoc* objections to the design). Still worse, the trend toward

¹This white paper can thus be seen as a complement to the recent report to the APSA Council on the future of association publications (Publications Planning Ad Hoc Committee 2014). For a broader vision of how social science journal practice might need to change to increase the credibility of published results, see Green, Humphreys, and Smith (N.d.).

²For more on how more transparent and credible research practices can increase trust in political science findings, see James N. Druckman's white paper in this task force report.

publishing mostly or entirely statistically significant results appears to be worsening over time (Fanelli 2012). What can we do? While we would like to think that science is self-correcting, the evidence suggests that numerous aspects of the research and publication process contribute to the proliferation of false positive results (Ioannidis 2012). I propose three reforms in this section intended to counter these tendencies: offering an option for authors to submit pre-accepted articles, exploring the feasibility of results-blind peer review, and conducting post-publication replication audits of a random subset of articles.

Pre-accepted articles

One response to concerns about publication bias and replication failures is the preregistration of experimental trials – a practice that is mandated in some areas of medicine and is beginning to be done voluntarily by some social science researchers conducting field experiments (particularly in development economics) as well as isolated cases with as-yet unobserved or uncollected observational data (e.g., Monogan 2013; Grossman and Pierskalla N.d.). An emerging cross-disciplinary movement of researchers argues for preregistration because it forces authors to publicly disclose their hypotheses and analysis plan before data have been collected, which should reduce the risk of spurious results (e.g., Gerber and Malhotra 2008a; Humphreys, de la Sierra, and van der Windt 2013; Wagenmakers et al. 2012; Monogan 2013; Miguel et al. 2014). One of the best examples of this practice to date is the Oregon Health Insurance Experiment. The authors publicly archived their analysis plan before data were available (Finkelstein and Baicker N.d.) and have explicitly labeled all unplanned analyses in published studies resulting from the experiment (Finkelstein et al. 2012; Baicker et al. 2013; Taubman et al. 2014 see also Casey, Glennerster, and Miguel 2012). In political science, Columbia’s Macartan Humphreys, Raul Sanchez de la Sierra, and Peter van der Windt have likewise proposed comprehensive but nonbinding registration for experiments (2013).

Unfortunately, preregistration alone will not solve the problem of publication bias. The comprehensive report format limits authors’ ability to produce the statistically significant findings that reviewers and editors demand and may lead authors to opt out of registration or to shelve non-significant findings. As a result, authors have little incentive to engage in the practice unless it offers an appealing route to publication in prestigious journals.³ In addition, authors may be tempted to deviate from preregistered analysis plans when they are not linked to journal publication practices. But most fundamentally, if studies are more likely to be published when they report statistically significant results, then publication bias is still likely to ensue even when preregistration is commonly practiced, as studies have frequently found in examining clinical trial results (e.g., Dwan et al. 2013).⁴

³Another approach — though one that is outside the scope of this article — would be to use funding agencies for leverage. As Said (2012) notes, funding agencies like the National Science Foundation could help change the incentives facing researchers by giving priority to scientists who publish in “outcome-unbiased” journals.

⁴See Humphreys, de la Sierra, and van der Windt (2013) for a stylized model of how the distribution of results across categories of publications might change under a voluntary, non-binding preregistration scheme.

A better practice would be for journals to offer authors an option in which articles with prespecified analysis plans would be accepted in principle *before* the study was conducted. By offering this voluntary option to authors of planned confirmatory studies, journals would create a positive incentive for preregistration that would avoid file drawer bias without imposing mandates or constraints that have raised concerns among some scholars (e.g., Laitin 2013; Publications Planning Ad Hoc Committee 2014). While this approach, which is known as Registered Reports, may seem radical, the format is now being offered by a number of journals, including *AIMS Neuroscience* (Chambers et al. 2014), *Cortex* (2013), *Perspectives on Psychological Science* (N.d.), and *Social Psychology* (Nosek and Lakens 2014).⁵ Most notably, a planned special issue of *Comparative Political Studies* will offer preaccepted articles for the first time in the discipline (Findley et al. 2014).

In this format, the article’s introduction, theory, and methods section would be written in advance. Articles would be accepted in principle after a first-stage review to prevent editors and reviewers from backing out of publishing an article due to a null or mixed finding. The authors would then conduct the study and populate the results section based on a prespecified analysis plan. Post-hoc analysis and interpretation would be allowed — scientific discovery is important! — but would have to be explicitly labeled as exploratory. The article would then be reviewed in a second round to ensure that the theory and hypotheses had not been changed, that any results that deviate from the prespecified plan are identified appropriately, and that the analysis of the results was conducted professionally.⁶ Rejecting the article on the basis of a null or mixed result would be prohibited in principle.⁷

While the pre-accepted article format is clearly most useful for surveys and experiments, the studies in question could include other types of observational data. The problem, however, is that pre-acceptance of these types of observational studies would require editors to trust authors’ self-reports that the data had not been previously collected or analyzed, creating an incentive problem of the type that preacceptance is intended to alleviate.⁸ An alternate approach would be to encourage authors of observational studies to submit articles that offer out-of-sample predictions (e.g., Monogan 2013). For instance, scholars who study democratic stability could offer forecasts of future irregular regime changes or coups over a prespecified period that would be evaluated after a specified interval of, say, 2–4 years (e.g., Beger 2014; Ulfelder 2014).

⁵Full disclosure: The author is a member of an ad hoc committee promoting the use of the Registered Reports format (<https://osf.io/8mpji/>).

⁶One potential variant on this design would be to include exploratory results as part of the first-round submission and then have a confirmatory analysis that is handled under the prespecification guidelines discussed above. Alternatively, journals could allow authors to conduct followup rounds of prespecified experiments or study waves if more than one set of results were required for the article, though this option could increase the burden on reviewers and editors.

⁷It might be possible to design a similar approach for qualitative research designs, but such a format is beyond the scope of this white paper or the expertise of the author.

⁸By contrast, experimental or survey data collection can be more easily documented as having been conducted *after* the design was accepted.

Despite these benefits, it is still possible that relatively few authors would use a preaccepted article format. To address this concern, the *American Political Science Review* (APSR) and other political science journals should consider offering incentives for authors to use the format by publishing preaccepted submissions as the lead articles in each issue and fast-tracking them in the editorial and peer review process to the extent possible. It might even be possible to raise funds to offer to authors whose designs are accepted in this format as *Perspectives on Psychological Science* now does (N.d.) or for the APSR to partner with Time-Sharing Experiments in the Social Sciences (TESS) to pre-accept articles that would then be fielded for free as part of a TESS omnibus survey.

Finally, it is worth noting the benefits that a shift toward a preaccepted article format could have for research and publishing practices in the discipline as a whole. The format would create healthy pressure on authors, editors, and reviewers to (a) identify hypotheses that are substantively important and developed carefully in relation to the previous literature; (b) answer research questions for which results would be worth publishing even if they are not surprising or counter-intuitive; (c) keep articles short given the likelihood of null or mixed results;⁹ and (d) make sure studies have sufficient statistical power to detect the hypothesized effect most of the time. In this way, it could have positive spillover effects even if the format is infrequently used.¹⁰

Results-blind peer review

One alternative approach to countering publication bias that is less closely linked to the experimental approach would be to offer “result-blind peer review” for quantitative studies of any sort (e.g., Glymour and Kawachi 2005; Hanson 2010; Greve, Bröder, and Erdfelder 2013; Smulders 2013; Green, Humphreys, and Smith N.d.). In this approach, reviewers would have to assess the theory and research design of a manuscript and make an initial decision without access to the statistical findings, which would be initially withheld. Such a decision would hopefully at least partly constrain the *post hoc* reasoning that helps produce publication bias against null results. Editors and reviewers would then assess the merits of a manuscript in a second round when the results were unblinded.

This approach faces significant implementation challenges. First, a small pilot study of this approach by Sridharan and Greenland (2009) at *Archives of Internal Medicine* found that there is reason for concern about publication bias even with a two-step review process. “[O]ver 7% of positive articles benefited from editors changing their minds between steps 1 and 2 of the alternate review,” they found, but “this never occurred with the negative studies.” Journals would have to be careful to try to avoid allowing publication bias to creep back in after the results were known to reviewers and editors. In addition, some evaluations might not be truly results-blind if editors and reviewers have previously read the paper or seen it at a conference (a particular problem in social science, which has long

⁹The shift toward shorter articles and online publishing may help to alleviate the space pressures on journals that article pre-acceptance could create.

¹⁰It is of course not possible to address every concern or objection to the format in the space available in this white paper. See Chambers et al. (2014) for responses to frequent objections to the format that have been raised by critics.

publication lags). Finally, offering this format as an option rather than a requirement could create an adverse selection problem in which results-blind submissions would mostly consist of articles with null or mixed results, allowing reviewers to infer that the authors' hypotheses were likely not supported. This inference could create potential biases against the authors' theory and design in first-stage reviews, allowing publication bias to creep back in, while also reducing the perceived value or status of articles published in the format. To avoid an adverse selection, journals might therefore consider randomizing some subset of manuscripts into this process or making it mandatory for certain special issues or sections (e.g., Findley et al. 2014).

Verifying replication data and code

Finally, every APSA journal should follow the practice of journals like the *American Journal of Political Science* (N.d.) and *American Economic Review* (AER; N.d.), which require authors of quantitative studies to provide a full replication archive before publication. This change would be consistent with the standards of data access that are now part of the association's code of ethics (APSA Committee on Professional Ethics, Rights, and Freedoms 2012; see Lupia and Elman 2014). As King (1995) and many others have argued, replication increases the likelihood of detecting errors and building cumulative scientific knowledge, helping to improve the rigor and value of published results in our field. Journals should therefore publish more replications of major results, though how to do so is a contested and difficult issue (see Ishiyama 2014 for a discussion of several possible approaches). However, *APSR* and other journals could go even further to encourage careful research practice by hiring qualified graduate students to ensure that the results of accepted articles can be reproduced from submitted replication files before publication like *Quarterly Journal of Political Science* (N.d.). By institutionalizing a practice that is typically employed only on ad hoc basis in class assignments (King 2006) and demonstration projects (Dewald, Thursby, and Anderson 1986), a mandatory journal replication policy would improve the incentives for scholars to engage in careful and systematic research practices. For instance, one in five articles examined from the 2006–2008 period in *AER* did not fully satisfy the requirement that results be reproducible from submitted data and code, leading the journal to require review by contracted grad students (Glandon 2010).

Improving political science review processes

The output rate and knowledge value of political science research depends significantly on the peer review and editorial processes within journals, which are intended to identify the most important manuscripts and improve their quality in a reasonably timely manner before publication. However, the incentives to provide rapid and rigorous reviews are weak, especially for the reviewers in greatest demand, who are frequently overburdened. I propose two strategies to improve the status quo: a frequent flier system for journals that would reward scholars who return quality reviews by the specified deadline and an option to allow authors to forward their manuscript and referee reports from the *APSR* to appropriate field journals, bypassing an additional round of review that may be unnecessary.

An additional concern is that unconscious biases may creep into editorial decisionmaking, distorting manuscript evaluations (the reason our discipline almost exclusively uses double-blind review). I therefore suggest the implementation of “triple-blind” reviewing, which would shield the identity of an article’s author(s) from editors as well as reviewers and thereby maximize the extent to which decisions are made based purely on merit.

A frequent flier system for journals

Journals depend on the free labor provided by academics in the peer review process. Reviewing is a largely thankless task whose burden falls disproportionately on prominent and public-minded scholars, who receive little credit for the work that they do. Academics therefore often prioritize other tasks over the reviews they have committed to provide. As a result, manuscripts are often stuck in review limbo for months, slowing the publication process and stalling the production of knowledge that our field can share with the public.

One idea is to develop a points system for *APSR* or all the APSA journals (as well as participating section journals) that is analogous to frequent flier miles. Each review would earn a scholar a certain number of points with bonuses awarded by editors for especially timely or high quality reviews. The author could then cash in those points when they submit to that journal in order to request a rapid review of their own manuscript.¹¹

The journal would in turn offer those points to reviewers who review the manuscript quickly, helping to speed it through the process. This system would not be useful for reviewers who do not submit often to the journal in question, but for reviewers and authors who frequently interact with a journal like the *APSR* over a period of decades, it could help provide greater incentives for rapid and thoughtful reviewing.¹²

Referee report rolldowns to section journals

One impediment to the rapid dissemination of knowledge from political science is the serial nature of the journal submission process. In many cases, it can take 6–12 months or more to receive reviews back from an initial submission to a journal. If that submission is declined, authors must resubmit to a new journal and start the process over again, which can frequently result in duplication of effort by reviewers and unnecessary delays for authors. One way to improve the efficiency of this process would be for APSA to offer

¹¹See Diana Mutz’s white paper in this task force report for more on how such a model might work if it extended across journals. My intention here is show how such a model could be feasible even if pursued independently by a single journal.

¹²APSA might also consider offering prizes for reviewers whose contributions are especially insightful and submitted by the specified deadline, which could help increase the incentives for timely, high quality reviews. Currently, no incentive exists like in medicine, where physicians can earn continuing medical education credits for reviews that are certified by journal editors as meeting expected standards of timeliness, quality, and detail (De Gregory 2004). For an example from economics of how to recognize a select group of referees, see the *American Economic Review* award for “Excellence in Refereeing” (http://www.aeaweb.org/aer/exc_ref.pdf).

authors the option of having their submissions to the *APSR* and the resulting reviews to be forwarded to other APSA or affiliated section journals. Many publications are declined at highly selective journals like the *APSR* with reviews that indicate that the paper should move forward to publication at a field journal. By allowing authors the option to redirect the manuscript and reviews to such a journal directly (a process that should be possible in online publishing systems), authors would be able to move their papers toward print more quickly while reducing the burden on scholars who often review a manuscript two or more times during the journal submission process.¹³ The American Economic Association has already implemented such a system for the *American Economic Review* and its affiliated *American Economic Journals* field journals. Authors are provided the option to forward referee reports and correspondence from the *AER* to *AEJ* editors (see, e.g., American Economic Journal: Applied Economics N.d.). While APSA does not publish field journals, many of its affiliated sections do (e.g., *Legislative Studies Quarterly*, *Political Analysis*, and *Political Behavior*). The association should consider proposing such a rolldown policy for those journals and explore partnerships with other field journals.¹⁴

Triple-blind reviewing

A number of studies have documented potential biases in evaluating scientific articles and abstracts as well as grant, fellowship, or job applications. For instance, scientific abstract submissions were found to be evaluated more favorably when originating in the U.S. or elite universities than those from people with other backgrounds (Ross et al. 2006). In addition, several studies have found that submissions and applications from women are viewed less favorably than those of men (Wennerås and Wold 1997; Steinpreis, Anders, and Ritzke 1999; Moss-Racusin et al. 2012), though these results have generally not replicated in larger samples (Ceci and Williams 2011). Concerns about potential biases like these are one reason for the dominance of double-blind review procedures among journals in our field, but they apply no less to editors than referees. Why would the author's identity be material to a decision about the merits of a manuscript?

Given these concerns, all APSA journals should conceal the identity of the author(s) from journal editors just as they do for referees — a so-called “triple-blind” review process. Of course, editors may have previously encountered the manuscript or could use Google to try to determine the author's identity (just as is true for reviewers). Some have suggested giving in to these pressures by possibly moving to a single-blind review system (e.g., Publications Planning Ad Hoc Committee 2014). I would argue that we should instead seek to minimize potential biases in editorial and review decisions to the extent possible.

¹³Ideally, referee identities would remain blinded from the author but made available to the editors at the rolldown journal, allowing them to assess the expertise of the reviewer.

¹⁴It is possible that such a mechanism could create incentives for rolldowns of low quality manuscripts, but presumably authors would be reluctant to forward negative reviews. If the pool is sufficiently skewed, however, a modest submission fee could be required that would be refunded upon publication at a rolldown journal. Journals might also consider empowering editors to reject rolldown submissions without review at a higher rate than regular articles.

This procedure is already used by the philosophy journals *Mind* (N.d.) and *Ethics* (N.d.) as well as *American Business Law Journal* (Cahoy 2010) and could easily be implemented in online manuscript review systems.¹⁵

By helping to minimize any inadvertent or unconscious biases in editorial decision-making, triple-blind review could improve the quality of publication decisions and increase confidence in the scientific integrity of the process among association members and disciplinary stakeholders.

Conclusion

Despite growing concerns about publication bias and increased enthusiasm for preregistration and replication, the incentives for change in current research practices remain quite weak. Unless leading journals like the *APSR* take steps like those described above to address publication bias, the status quo is likely to remain in place. Similarly, concerns about reviewer burdens and long lag times before publication have persisted for years without a solution due to a lack of institutional changes.

The Association has the opportunity to institute best practices in journal publishing that would match or exceed existing standards across the social and natural sciences. These changes would enhance the credibility of political science research, accelerate the pace at which it is evaluated, and increase its evidentiary value.

The changes I propose are thus complementary to the other initiatives to promote public engagement that are described in this task force report. If we hope to engage the public successfully, it is essential that we increase trust in political science research. Conversely, if we are successful at attracting attention to our scholarship from government officials, journalists, and civil society organizations, it is vital that the research that we produce is as timely and credible as possible. We should not miss this opportunity to make our journal practices consistent with the highest aspirations of our discipline.

References

American Economic Journal: Applied Economics. N.d. “Editorial Policy.” Downloaded February 7, 2014 from <https://www.aeaweb.org/aej/app/edpolicy.php>.

American Economic Review. N.d. “The American Economic Review: Data Availability Policy.” Downloaded February 6, 2014 from <http://www.aeaweb.org/aer/data.php>.

American Journal of Political Science. N.d. “Guidelines for accepted articles.” Downloaded February 6, 2014 from <http://ajps.org/guidelines-for-accepted-articles/>.

¹⁵It would require one additional step, however — a graduate student assistant would be needed to identify any conflicts of interest between the referees selected by the editor and the author before any review requests were made.

- APSA Committee on Professional Ethics, Rights, and Freedoms. 2012. "A Guide to Professional Ethics in Political Science." Second Edition. Downloaded February 7, 2014 from <http://www.apsanet.org/media/PDFs/ethicsguideweb.pdf>.
- Baicker, Katherine, Sarah L. Taubman, Heidi L. Allen, Mira Bernstein, Jonathan H. Gruber, Joseph P. Newhouse, Eric C. Schneider, Bill J. Wright, Alan M. Zaslavsky, and Amy N. Finkelstein. 2013. "The Oregon Experiment — Effects of Medicaid on Clinical Outcomes." *New England Journal of Medicine* 368 (18): 1713–1722.
- Beger, Andreas. 2014. "The coup in Thailand and progress in forecasting." *Predictive Heuristics*. May 22, 2014. Downloaded June 13, 2014 from <http://predictiveheuristics.com/2014/05/22/the-coup-in-thailand-and-progress-in-forecasting/>.
- Button, Katherine S., John P. A. Ioannidis, Claire Mokrysz, Brian A Nosek, Jonathan Flint, Emma S.J. Robinson, and Marcus R. Munafò. 2013. "Power failure: why small sample size undermines the reliability of neuroscience." *Nature Reviews Neuroscience* 14: 365–376.
- Cahoy, Daniel R. 2010. "Editor's Corner: Assembling a Special Issue on Law as a Source of Strategic Advantage." *American Business Law Journal* 47 (4): v–viii.
- Casey, Katherine, Rachel Glennerster, and Edward Miguel. 2012. "Reshaping Institutions: Evidence on Aid Impacts Using a Preanalysis Plan*." *Quarterly Journal of Economics* 127 (4): 1755–1812.
- Ceci, Stephen J., and Wendy M. Williams. 2011. "Understanding current causes of women's underrepresentation in science." *Proceedings of the National Academy of Sciences* 108 (8): 3157–3162.
- Chambers, Christopher D. 2013. "Registered Reports: A new publishing initiative at Cortex." *Cortex* 49 (3): 609–610.
- Chambers, Christopher D., Eva Feredoes, Suresh D. Muthukumaraswamy, and Peter J. Etchells. 2014. "Instead of "playing the game" it is time to change the rules: Registered Reports at AIMS Neuroscience and beyond." *AIMS Neuroscience* 1 (1): 4–17.
- De Gregory, Jamie. 2004. "Medical journals start granting CME credit for peer review." *Science Editor* 27 (6): 190–191.
- Dewald, William G., Jerry G. Thursby, and Richard G. Anderson. 1986. "Replication in Empirical Economics: The Journal of Money, Credit and Banking Project." *American Economic Review* 76 (4): 587–603.
- Donnellan, M. Brent, Richard E. Lucas, and Joseph Cesario. N.d. "On the Association between Loneliness and Bathing Habits: Nine Replications of Bargh and Shalev (2012) Study." Forthcoming, *Emotion*.

- Doucouliaagos, Chris. 2005. "Publication bias in the economic freedom and economic growth literature." *Journal of Economic Surveys* 19 (3): 367–387.
- Dwan, Kerry, Carrol Gamble, Paula R. Williamson, and Jamie J. Kirkham. 2013. "Systematic review of the empirical evidence of study publication bias and outcome reporting bias—An updated review." *PLOS One* 8 (7): e66844.
- Esarey, Justin, and Ahra Wu. N.d. "The fault in our stars: Measuring and correcting significance bias in Political Science." Unpublished manuscript.
- Ethics. N.d. "What Should I Expect Once I've Submitted?" Downloaded February 6, 2014 from <http://www.press.uchicago.edu/journals/et/afterSubmission.html?journal=et>.
- Fanelli, Daniele. 2012. "Negative results are disappearing from most disciplines and countries." *Scientometrics* 90 (3): 891–904.
- Ferguson, Christopher J., and Moritz Heene. 2012. "A vast graveyard of undead theories publication bias and psychological science's aversion to the null." *Perspectives on Psychological Science* 7 (6): 555–561.
- Findley, Michael, Nathan Jensen, Edmund Malesky, and Thomas Pepinsky. 2014. "Call for Papers: Special Issue of *Comparative Political Studies* on Research Transparency in the Social Sciences." Downloaded June 11, 2014 from <http://www.ipdutexas.org/cps-transparency-special-issue.html>.
- Finkelstein, Amy, and Katherine Baicker. N.d. "Oregon Health Insurance Experiment: Analysis Plans." Downloaded February 5, 2014 from <http://www.nber.org/oregon/documents.html>.
- Finkelstein, Amy, Sarah Taubman, Bill Wright, Mira Bernstein, Jonathan Gruber, Joseph P Newhouse, Heidi Allen, Katherine Baicker, and Oregon Health Study Group. 2012. "The Oregon Health Insurance Experiment: Evidence from the First Year*." *Quarterly Journal of Economics* 127 (3): 1057–1106.
- Gerber, Alan, and Neil Malhotra. 2008a. "Do Statistical Reporting Standards Affect What Is Published? Publication Bias in Two Leading Political Science Journals." *Quarterly Journal of Political Science* 3: 313–326.
- Gerber, Alan S, and Neil Malhotra. 2008b. "Publication Bias in Empirical Sociological Research: Do Arbitrary Significance Levels Distort Published Results?" *Sociological Methods & Research* 37 (1): 3–30.
- Gerber, Alan S., Neil Malhotra, Conor M. Dowling, and David Doherty. 2010. "Publication bias in two political behavior literatures." *American Politics Research* 38 (4): 591–613.
- Glandon, Philip. 2010. "Report on the *American Economic Review* data availability compliance project." Unpublished manuscript.

- Glymour, M Maria, and Ichiro Kawachi. 2005. "Review of publication bias in studies on publication bias: Here's a proposal for editors that may help reduce publication bias." *BMJ: British Medical Journal* 331 (7517): 638.
- Green, Don, Macartan Humphreys, and Jenny Smith. N.d. "Read it, understand it, believe it, use it: Principles and proposals for a more credible research publication." Unpublished manuscript.
- Greve, Werner, Arndt Bröder, and Edgar Erdfelder. 2013. "Result-Blind Peer Reviews and Editorial Decisions." *European Psychologist* 18 (4): 1–9.
- Grossman, Guy, and Jan H. Pierskalla. N.d. "The Effects of Administrative Unit Proliferation on Service Delivery." Unpublished manuscript.
- Hanson, Robin. 2010. "Result Blind Review." November 6, 2010. Downloaded January 28, 2014 from <http://www.overcomingbias.com/2010/11/results-blind-peer-review.html>.
- Humphreys, Macartan, Raul Sanchez de la Sierra, and Peter van der Windt. 2013. "Fishing, commitment, and communication: A proposal for comprehensive nonbinding research registration." *Political Analysis* 21 (1): 1–20.
- Ioannidis, John P.A. 2005. "Why most published research findings are false." *PLoS medicine* 2 (8): e124.
- Ioannidis, John P.A. 2012. "Why science is not necessarily self-correcting." *Perspectives on Psychological Science* 7 (6): 645–654.
- Ishiyama, John. 2014. "Replication, Research Transparency, and Journal Publications: Individualism, Community Models, and the Future of Replication Studies." *PS: Political Science & Politics* 47 (01): 78–83.
- King, Gary. 1995. "Replication, replication." *PS: Political Science and Politics* 28 (3): 444–452.
- King, Gary. 2006. "Publication, publication." *PS* 39 (1): 119–125.
- Laitin, David D. 2013. "Fisheries Management." *Political Analysis* 21 (1): 42–47.
- Lupia, Arthur, and Colin Elman. 2014. "Openness in Political Science: Data Access and Research Transparency." *PS: Political Science & Politics* 47 (01): 19–42.
- Masicampo, E.J., and Daniel R. Lalande. 2012. "A peculiar prevalence of p values just below .05." *Quarterly Journal of Experimental Psychology* 65 (11): 2271–2279.
- Miguel, E., C. Camerer, K. Casey, J. Cohen, K. M. Esterling, A. Gerber, R. Glennerster, D. P. Green, M. Humphreys, G. Imbens, D. Laitin, T. Madon, L. Nelson, B. A. Nosek, M. Petersen, R. Sedlmayr, J. P. Simmons, U. Simonsohn, and M. Van der Laan. 2014. "Promoting Transparency in Social Science Research." *Science* 343 (6166): 30–31.

- Mind. N.d. "Review Procedure." Downloaded February 6, 2014 from http://www.oxfordjournals.org/our_journals/mind/review_procedure.html.
- Monogan, James E. 2013. "A case for registering studies of political outcomes: An application in the 2010 House elections." *Political Analysis* 21 (1): 21–37.
- Montgomery, Jacob, and Brendan Nyhan. 2010. "Bayesian Model Averaging: Theoretical developments and practical applications." *Political Analysis* 18 (2): 245–270.
- Moss-Racusin, Corinne A., John F. Dovidio, Victoria L. Brescoll, Mark J. Graham, and Jo Handelsman. 2012. "Science faculty's subtle gender biases favor male students." *Proceedings of the National Academy of Sciences* 109 (41): 16474–16479.
- Nosek, Brian A., and Dani'el Lakens. 2014. "Registered Reports: A Method to Increase the Credibility of Published Results." *Social Psychology* 45 (3): 137–141.
- Nyhan, Brendan. 2012a. "Academic reforms: A four-part proposal." April 16, 2012. Downloaded February 6, 2014 from <http://www.brendan-nyhan.com/blog/2012/04/academic-reforms-a-four-part-proposal.html>.
- Nyhan, Brendan. 2012b. "More on pre-accepted academic articles." April 27, 2012. Downloaded February 6, 2014 from <http://www.brendan-nyhan.com/blog/2012/04/more-on-pre-accepted-academic-articles.html>.
- Nyhan, Brendan, John Sides, and Joshua A. Tucker. N.d. "APSA as amplifier: How to encourage and promote public voices within political science." Unpublished manuscript.
- of Political Science, Quarterly Journal. N.d. "Author Instructions." Downloaded June 13, 2014 from <http://nowpublishers.com/journals/QJPS/author-instructions>.
- Perspectives on Psychological Science*. N.d. "Registered Replication Reports." Downloaded February 5, 2014 from <http://www.psychologicalscience.org/index.php/replication>.
- Publications Planning Ad Hoc Committee. 2014. "Report to APSA Council." *PS* 47 (01): 246–256.
- Ritchie, Stuart J., Richard Wiseman, and Christopher C. French. 2012. "Failing the future: Three unsuccessful attempts to replicate Bem's 'Retroactive facilitation of recall' effect." *PLOS One* 7 (3): e33423.
- Ross, Joseph S., Cary P. Gross, Mayur M. Desai, Yuling Hong, Augustus O. Grant, Stephen R. Daniels, Vladimir C. Hachinski, Raymond J. Gibbons, Timothy J. Gardner, and Harlan M. Krumholz. 2006. "Effect of blinded peer review on abstract acceptance." *JAMA* 295 (14): 1675–1680.
- Said, Chris. 2012. "It's the incentive structure, people!" Downloaded February 5, 2014 from <http://filedrawer.wordpress.com/2012/04/17/>.

- Simmons, Joseph P., Leif D. Nelson, and Uri Simonsohn. 2011. "False-positive psychology undisclosed flexibility in data collection and analysis allows presenting anything as significant." *Psychological Science* 22 (11): 1359–1366.
- Simonsohn, Uri, Leif Nelson, and Joseph Simmons. 2013. "P-curve: A key to the file drawer." *Journal of Experimental Psychology: General*.
- Smulders, Yvo M. 2013. "A two-step manuscript submission process can reduce publication bias." *Journal of Clinical Epidemiology* 66 (9): 946–947.
- Sridharan, Lakshmi, and Philip Greenland. 2009. "Editorial policies and publication bias: The importance of negative studies." *Archives of Internal Medicine* 169 (11): 1022–1023.
- Steinpreis, Rhea E., Katie A. Anders, and Dawn Ritzke. 1999. "The impact of gender on the review of the curricula vitae of job applicants and tenure candidates: A national empirical study." *Sex Roles* 41 (7/8): 509–528.
- Taubman, Sarah L., Heidi L. Allen, Bill J. Wright, Katherine Baicker, and Amy N. Finkelstein. 2014. "Medicaid Increases Emergency-Department Use: Evidence from Oregon's Health Insurance Experiment." *Science* 343 (6168): 263–268.
- Ulfelder, Jay. 2014. "Coup Forecasts for 2014." Dart-Throwing Chimp. January 25, 2014. Downloaded June 13 from <http://dartthrowingchimp.wordpress.com/2014/01/25/coup-forecasts-for-2014/>.
- Wagenmakers, Eric-Jan, Ruud Wetzels, Denny Borsboom, Han L.J. van der Maas, and Rogier A. Kievit. 2012. "An agenda for purely confirmatory research." *Perspectives on Psychological Science* 7 (6): 632–638.
- Wennerås, Christine, and Agnes Wold. 1997. "Nepotism and sexism in peer-review." *Nature* 387: 341–343.