SECTION II: CHANGING INCENTIVES

Increasing the Credibility of Political Science Research: A Proposal for Journal Reforms

Brendan Nyhan, Dartmouth College

How can political science improve its standing in and contribution to public life? In addition to engaging more effectively in the national debate (see Nyhan, Sides, and Tucker 2015, this special issue), we also can strengthen the quality and timeliness of the scholarship that we hope to bring to a wider audience.

This article proposes two sets of reforms to APSA journal practices that are intended to achieve those goals. The first proposal seeks to counter publication bias and increase incentives for careful research practices that generate replicable results. The second proposal is designed to improve the timeliness and quality of reviewer and editorial decisions during the review process.

If implemented carefully, these reforms can make the publication process more rigorous and efficient, thereby enhancing the value of the findings we produce, which would be more timely and credible and therefore more valuable to the external audiences we want 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.

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 (Esarey and Wu n.d.; Gerber and Malhotra 2008a; Gerber et al. 2010), sociology (Gerber and Malhotra 2008b), economics (Doucouliagos 2005), and psychology (Ferguson and Heene 2012; Masicampo and Lalande 2012), which results in published literatures that contain too many narrowly statistically significant findings to be attributable to chance (i.e., heaping of p-values at slightly less than p < 0.05). Too often, such findings are susceptible 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 2015; Ritchie, Wiseman, and French 2012). By contrast, null findings typically are 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).

Although some fraud may occur, the problem is more likely to be one of self-deception—that is, as authors, reviewers, and editors, we are simply too practiced in rationalizing choices that produce the results we want, including “p-hacking” (Simmons, Nelson, and Simonsohn 2011; Simonsohn, Nelson, and Simmons 2013) and the use of low-power studies (Button et al. 2013). Moreover, our acceptance of these norms often leads to reviewer insistence on statistical significance or to rejections of articles that fail to reach this threshold based on other rationales (e.g., post-hoc objections to the research design). Still worse, the trend toward publishing mostly or entirely statistically significant results appears to be worsening over time (Fanelli 2012). What can we do? It is tempting to think that science is self-correcting, but the evidence suggests that numerous aspects of the research and publication process contribute to the proliferation of false-positive results (Ioannidis 2012) and need to be changed. In this section, I propose three reforms intended to counter these tendencies: (1) offering an option for authors to submit pre-accepted articles, (2) exploring the feasibility of results-blind peer review, and (3) conducting postpublication replication audits of a random subset of articles.

Preaccepted 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 happen voluntarily by social science researchers conducting field experiments (particularly in development economics) or studying unobserved or uncollected observational data (e.g., Grossman and Pierskall a n.d.; Monogan 2013). An emerging cross-disciplinary movement of researchers argues for preregistration because it forces authors to publicly disclose their hypotheses and analysis plans 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; Miguel et al. 2014; Monogan 2013; Wagenmakers et al. 2012). 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 (Finckelstein and Baicker n.d.) and explicitly labeled all unplanned analyses in published studies resulting from the experiment (Baicker et al. 2013; Finckelstein et al. 2012; Taubman et al. 2014; see also Casey, Glennerster, and Miguel 2012). In political science, Humphreys, de la Sierra, and van der Windt (2013) likewise proposed comprehensive but non-binding registration for experiments.

Unfortunately, preregistration alone will not solve the problem of publication bias. The format limits authors’ ability to produce the statistically significant findings that reviewers and editors demand, which may lead authors to opt out of registration or to shelve nonsignificant findings. As a result, authors have little incentive to engage in the practice unless it offers an appealing route to publication in prestigious journals. In addition, scholars may be tempted to deviate from preregistered analysis plans when those plans are not linked to journal publication practices. Most fundamental, however, if studies are more apt 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 frequently have found in examining results from clinical trials (see, e.g., Dwan et al. 2013).4

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 of planned confirmatory studies, journals would create a positive incentive for preregistration that would minimize file-drawer bias and avoid the mandates and constraints to which some scholars object (e.g., Laitin 2013; Publications Planning Ad Hoc Committee 2014). Although this approach, which is known as Registered Reports, may seem radical, the format is currently being offered by a number of journals, including AIMS Neuroscience (Chambers et al. 2014), Cortex (Chambers 2013), Perspectives on Psychological Science (n.d.), and Social Psychology (Nosek and Lakens 2014).5 Most notable, a planned special issue of Comparative Political Studies will offer preaccepted articles for the first time in the discipline (Findley et al. 2014).

Using this format, an article’s introduction, theory, and methods sections 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 then would 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 must be explicitly labeled as exploratory. The article then would 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 results was conducted professionally. Rejecting the article on the basis of a null or mixed result would be prohibited in principle.6

Although the preaccepted article format clearly is most useful for surveys and experiments, the studies in question could include other types of observational data. The problem, however, is that preacceptance of observational studies would require editors to trust authors’ self-reports that the data had not been collected or analyzed previously—which creates 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 throughout a prespecified period that would be evaluated after a specified interval of perhaps two to four years (e.g., Beger 2014; Ulfelder 2014).

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 lead articles in each issue and—to the extent possible—fast-tracking them in the editorial and peer-review process. It may be possible to raise funds to offer authors whose designs are accepted in this format, as currently done by Perspectives on Psychological Science (n.d.), or for the APSR to partner with Time-Sharing Experiments in the Social Sciences (TESS) to preaccept articles that then would 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 (1) identify hypotheses that are substantively important and developed carefully in relation to the previous literature; (2) answer research questions for which results would be worth publishing even if they are not surprising or countervuitive; (3) keep articles short given the likelihood of null or mixed results; and (4) ensure that

Using this format, an article’s introduction, theory, and methods sections 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.

Results-Blind Peer Review

An alternative approach to countering publication bias that is less closely linked to the experimental method is to offer “result-blind peer review” for quantitative studies of any type (e.g., Glymour and Kawachi 2005; Green, Humphreys, and Smith n.d.; Greve, Bröder, and Erdfelder 2013; Hanson 2010; Smulders 2013). In this approach, reviewers would assess the theory and research design of a manuscript and make an initial decision without access to the statistical findings, which initially would be withheld. It is hoped that such a decision would constrain (at least partly) the post-hoc reasoning that produces publication bias against null results. Editors and reviewers then would 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” (Sridharan and Greenland 2009). Journals would need to carefully avoid allowing publication bias to gradually return after the results were known to reviewers and editors. Second, some evaluations may not be truly results-blind if editors and reviewers previously read the paper or heard it presented at a conference (a particular problem in social science, which has long publication lags). Third, offering this format as an option rather than as a requirement could create an adverse selection problem in which results-blind submissions would consist mostly of articles with null or mixed results, thereby allowing reviewers to infer that
Section II: Increasing the Credibility of Political Science Research: A Proposal for Journal Reforms

By institutionalizing

Increasing the Credibility of Political Science Research: A Proposal for Journal Reforms

By allowing authors the option to redirect the manuscript and reviews to such a journal (a process that should be possible in online publishing systems), they could move toward publication more quickly while reducing the reviewing burden on other scholars during the submission process.

Verifying Replication Data and Code

Every APSA journal should follow the practice of journals including 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 currently part of the APSA’s code of ethics (APSA Committee on Professional Ethics, Rights, and Freedoms 2012; see also Lupia and Elman 2014). As King (1995) and many others argued, replication increases the likelihood of detecting errors and building cumulative scientific knowledge, which improves the rigor and value of published results in our field. Journals therefore should publish more replications of major results—although how to do so is a contested and difficult issue (for a discussion of several possible approaches, see Ishiyama 2014). However, APSR and other journals could further 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, as done at the Quarterly Journal of Political Science (n.d.). By institutionalizing a practice that typically is used only on an ad-hoc basis in class assignments (King 2006) and demonstration projects (Dewald, Thursby, and Anderson 1986), a mandatory journal-replication policy would improve the incentive for scholars to engage in careful and systematic research practices. For instance, one in five articles examined in 2006–2008 in AER did not fully satisfy the

By allowing authors the option to redirect the manuscript and reviews to such a journal (a process that should be possible in online publishing systems), they could move toward publication more quickly while reducing the reviewing burden on other scholars during the submission process.

requirement that results be reproducible from submitted data and code, which led the journal to require review by contracted graduate students (Glandon 2010).

IMPROVING POLITICAL SCIENCE REVIEW PROCESSES

The output rate and knowledge value of political science research significantly depends 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 in this section to improve the status quo: (1) a “frequent-flier” system for journals, which would reward scholars who return quality reviews by the specified deadline; and (2) an option to allow authors to forward their manuscripts and referee reports from the APSR to appropriate field journals, bypassing an additional and possibly unnecessary round of review.

Another concern is that unconscious biases may creep into editorial decision making, distorting manuscript evaluations (i.e., the reason that 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, thereby maximizing 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 generally thankless task that falls disproportionately on prominent and public-minded scholars who receive little credit for their work. Therefore, academics often prioritize other tasks before the reviews that they have committed to provide. As a result, manuscripts often are backlogged for months in “review limbo,” which slows the publication process and stalls the production of knowledge that our field can share with the public.

One idea is to develop a point system for APSR or for all of the APSA journals (as well as participating section journals) that is analogous to a frequent-flier program. Each review would earn a scholar a specific number of points, with bonuses awarded by editors for especially timely or high-quality reviews. To request a rapid review of their own manuscript, authors then could redeem those points when they submit to that particular journal.11

The journal, in turn, would offer points to those who review the manuscript quickly, thereby expediting the process. This system would not be useful for reviewers who do not make frequent submissions to a particular journal. However, for reviewers and

By allowing authors the option to redirect the manuscript and reviews to such a journal (a process that should be possible in online publishing systems), they could move toward publication more quickly while reducing the reviewing burden on other scholars during the submission process.

authors who frequently interact with a journal such as APSR during a period of decades, the system could provide greater incentives for rapid and thoughtful reviewing.12

Referee Report Roll-Downs 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 six to 12 months or more to receive reviews from an initial submission to a journal. If that submission is declined, the author must resubmit it to a new journal and restart the process, which frequently can result in duplication of effort by reviewers and unnecessary delays for authors. One way to improve the efficiency of this process is for the APSA to offer authors the option of forwarding their APSR submissions and the resulting reviews to other APSA or affiliated
section journals. Many submissions are declined at highly selective journals such as the APSR with reviews indicating that the manuscripts should move forward toward publication at field journals. By allowing authors the option to redirect the manuscript and reviews to such a journal (a process that should be possible in online publishing systems), they could move toward publication more quickly while reducing the reviewing burden on other scholars during the submission process.13 The American Economic Association implemented such a system for the AER and its affiliated American Economic Journal (AEJ) field journals. Authors are given the option to forward referee reports and correspondence from the AER to the AEJ editors (see, e.g., American Economic Journal: Applied Economics n.d.). Although the 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 roll-down policy for those journals and explore partnerships with other field journals.14

**Triple-Blind Reviewing**

Several studies have documented potential biases in evaluating scientific articles and abstracts as well as grant, fellowship, and job applications. For instance, scientific submissions were found to be evaluated more favorably when they originate in the United States or elite universities than those from people with a different background (Ross et al. 2006). In addition, several studies found that submissions and applications from women are viewed less favorably than those from men (Moss-Racusin et al. 2012; Steinpreis, Anders, and Ritzke 1999; Wennerås and Wold 1997), although these results generally have not been replicated in larger samples (Ceci and Williams 2011). Concerns about these potential biases are one reason for the dominance of double-blind review procedures among journals in our field, but they apply no less to editors than to referees. Why would an 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 a manuscript or can use Google to determine an author’s identity (which also is true for reviewers). Some have suggested bowing to these pressures by possibly moving to a single-blind review system (e.g., Publications Planning Ad Hoc Committee 2014). I argue that we instead should seek to minimize potential biases in editorial and review decisions. The triple-blind procedure already is used by the philosophy journals Mind (n.d.) and Ethics (n.d.) and the American Business Law Journal (Cahoy 2010) and easily could be implemented in online manuscript-review systems.15

By minimizing 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 APSA 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 weak. Unless leading journals such as the APSR take steps like those I describe to address publication bias, the status quo is likely to remain in place. Similarly, due to a lack of institutional changes, concerns about reviewer burdens and long lag times before publication have persisted for years without a solution.

The APSA 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 proposed in this article are thus complementary to other initiatives described in this task force report to promote public engagement. If we hope to successfully engage the public, it is essential that we increase trust in political science research. Conversely, if we are successful in attracting the attention of government officials, journalists, and civil-society organizations to our scholarship, it is vital that the research produced 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.

**ACKNOWLEDGMENTS**

This article was written for the American Political Science Association (APSA) Task Force on Public Engagement: Improving Public Perceptions of Political Science’s Value. I thank Deborah Brooks, John Carey, Brian Greenhill, Jeremy Horowitz, Arthur Lupia, and Ben Valentino for helpful comments and discussion.

The content of this article is based in part on Nyhan (2012a, 2012b). ■

**NOTES**

1. Thus, this article can 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 may need to change to increase the credibility of published results, see Green, Humphreys, and Smith (n.d.).

2. For information on how more transparent and credible research practices can increase trust in political science findings, see James N. Druckman’s article in this special issue.

3. Another approach—although one that is beyond the scope of this article—is to use funding agencies for leverage. As Said (2012) notes, grant makers such as the National Science Foundation could change the incentives facing...
Section II: Increasing the Credibility of Political Science Research: A Proposal for Journal Reforms

researchers by giving priority to scientists who publish in “outcome-based” journals. A review of publication bias could make the difference between a paper being published in a well-regarded journal and not being published at all. In this context, the use of publication bias as an indicator of the quality of research is essential. The proposed method could provide an objective measure of the quality of research, which could be used to evaluate the credibility of political science research.

4. See 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, nonbinding preregistration scheme.

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

6. A potential variant on this design is to include exploratory results as part of the first-round submission and then have a confirmatory analysis that is handled under the preregistration guidelines discussed herein. Alternatively, journals could allow authors to conduct follow-up rounds of preregistered analyses or study waves if more than one set of results were required for the article. However, this option could increase the burden on reviewers and editors.

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

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

9. The shift toward shorter articles and online publishing may alleviate the space restrictions on journals that article preacceptance could create.

10. Of course, it is not possible to address every concern about or objection to the format in this article. See Chambers et al. (2014) for responses to frequent objections to the format that have been raised by critics.

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

12. The APSA might also consider offering prizes for reviewers whose contributions are especially insightful and submitted by the specified deadline, which could increase the incentives for timely, high-quality reviews. Currently, no incentive exists for such work. The field in which 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 example, one from economists on how to recognize a select group of referees, see the American Economic Review award for “Excellence in Refereeing” (www.aeaweb.org/aer/exc_ref.pdf).

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

14. It is possible that such a mechanism could create incentives for roll-downs of low-quality manuscripts; presumably, however, 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 on publication at a roll-down journal.

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

REFERENCES


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.


The APSA Mentoring Program is an opportunity for all political scientists to network and share experience, advice and insights about career planning, scholarship, and a variety of professional issues such as:

- Selecting a dissertation committee
- Publishing and research
- Non-academic careers
- Life as junior faculty member
- Balancing family and work

The APSA mentoring database includes mentors with a wide variety of expertise and fields of interest. APSA is looking for additional mentors specifically in non-academic careers and international relations, and to mentor Ralph Bunche Scholars (RBSI) and/or APSA Minority Fellows (MFP).

Use the online submission form to become or request a mentor!

Sign up at www.apsanet.org/mentor.