Standard Operating Procedures: A Safety Net for Pre-Analysis Plans

Winston Lin, *Columbia University*
Donald P. Green, *Columbia University*

ABSTRACT Across the social sciences, growing concerns about research transparency have led to calls for pre-analysis plans (PAPs) that specify in advance how researchers intend to analyze the data they are about to gather. PAPs promote transparency and credibility by helping readers distinguish between exploratory and confirmatory analyses. However, PAPs are time-consuming to write and may fail to anticipate contingencies that arise in the course of data collection. This article proposes the use of "standard operating procedures" (SOPs)—default practices to guide decisions when issues arise that were not anticipated in the PAP. We offer an example of an SOP that can be adapted by other researchers seeking a safety net to support their PAPs.

FOREWORD BY ERIC S. DICKSON, EDITOR, *JOURNAL OF EXPERIMENTAL POLITICAL SCIENCE*

"Standard Operating Procedures: A Safety Net for Pre-Analysis Plans" addresses an issue of the highest importance in experimental political science. In this foreword, I provide a journal editor's perspective on why this piece is likely to have a significant impact on the way in which future research is carried out in our discipline.

In recent years, the reliability of published research in the social sciences has been subject to great scrutiny not only in academic circles but also in the popular press, most notably highlighted by recent findings that many well-known studies in social psychology do not replicate. A number of interrelated factors are implicated in this regrettable situation, including publication bias, "p-hacking," and "harking" (hypothesizing after the results are known). The quality and usefulness of the published body of social scientific research is very much in question, a matter all the more urgent in an era when funding for social science research is a red-hot political issue. The silver lining in this crisis is that some researchers and journal editors are starting to get serious about improving standards for and basic practices in social science research, and there is every prospect that our disciplines can emerge stronger, more reliable, and more worthy of respect as a result of increased thoughtfulness and institutional reforms.

One important trend has been a growing interest in, and use of, preregistration of experimental research designs, including detailed pre-analysis plans. Investigators who file preregistered designs and pre-analysis plans commit themselves to stating their hypotheses ex ante, preventing "harking" and clarifying the distinction between confirmatory and exploratory research. They also commit themselves to presenting the results of analysis methods that are preferred for ex ante theoretical reasons, limiting the ability of researchers to "fish" (consciously or unconsciously) for those analysis specifications that make their results appear (in a biased manner) more striking ex post. The use of preregistered designs and pre-analysis plans has been growing, and will doubtless continue to grow even more quickly over the next few years.

Around 20 academic journals in the human behavioral sciences, particularly in psychology and neuroscience, have taken the further step of opening a "registered reports" submission track, whereby research is accepted "in principle" for publication upon peer review of a detailed preregistered research design and pre-analysis plan, before data collection has taken place. This reform addresses publication bias, and commits both journal editors and authors to publish the results of well-designed and theoretically important research, regardless of whether these results are statistically significant. Within political science, Comparative Political Studies *has led the way, using results-free review in compiling its "Special Issue on Research Transparency in the Social Sciences." As editor of the* Journal of Experimental Political Science, *I will be bringing a*

Winston Lin *is a postdoctoral research scholar in political science at Columbia University. He can be reached at Linston@gmail.com.*
Donald P. Green *is a professor of political science at Columbia University. He can be reached at dpg2110@columbia.edu.*

doi:10.1017/S1049096516000810 © American Political Science Association, 2016
https://doi.org/10.1017/S1049096516000810 Published online by Cambridge University Press

permanent "registered reports" submission track to political science for the first time later in 2016.

The use of preregistration and pre-analysis plans sounds simple in the abstract. But in practice, researchers may not anticipate certain pitfalls that will occur when the research is actually implemented and the relevant data collected. Through no fault of their own, researchers may find themselves in situations where they face the challenge of analyzing data that was generated through a process not exactly corresponding to their preregistered design, and which as a result cannot sensibly be analyzed exactly according to the corresponding pre-analysis plan. To the extent that researchers then deal with the necessary deviations from preregistered procedures in an ad hoc manner, we are in a certain respect back where we started in the first place (though, of course, in not such a severe position)—researchers' discretion in methods of analysis means the possibility of conscious or unconscious "fishing," and means that submitted and published research will be biased accordingly.

This article's key insight is that, by precommitting to Standard Operating Procedures (SOPs) for dealing with a range of potentially anticipatable pitfalls in the implementation of experimental research, researchers can bind themselves to following analysis procedures that were ex ante justifiable on methodological or theoretical grounds, thus limiting the discretion they could exercise in choosing analysis methods with data in hand. This is crucial in my view for at least three important reasons. First, pitfalls are inevitable, even in projects carried out by the most skilled and careful researchers. The problem that SOPs intend to ameliorate is thus a widespread and important one. Second, the presence of SOPs reduces the cost of preparing a detailed preregistered design / pre-analysis plan, thus increasing the likelihood that researchers will use these important tools in the first place. This is so because certain off-the-shelf or previously settled protocols will be one-time "fixed costs" (not even necessarily borne by the researcher, if she adopts a publicly circulating set of SOPs) rather than "marginal costs" associated with every research project. And third, the existence of SOPs will make it easier for editors to implement registered reports submissions tracks, thereby aiding the struggle against publication bias and making it easier to incentivize creative, original research whose results are to be believed.

Of course, by their nature no set of SOPs will ever be fully complete. But I anticipate that the scope of publicly circulating SOPs will grow over time, eventually extending to the issue of on-the-fly adjustments to experimental designs, especially for field researchers operating in unpredictable and logistically challenging environments.

Social science research is at a crossroads. Through a series of ingenious insights, our research community has discovered how much scope there is radically to improve the reliability of published research. Following through on this potential is crucial for the future of our discipline. Few concrete contributions toward this end that I have encountered are as practically useful, and intellectually interesting, as the authors' advocacy for Standard Operating Procedures. ■

For decades, scholars have warned about biases in the empirical research literature due to "data fishing" and "file drawer" problems. The fishing problem is that there are many possible ways to analyze any data set, and decisions about how to analyze the data and which analyses to report may be influenced by their effects on the results (Humphreys, Sanchez de la Sierra, and van der Windt 2013). Fishing can take many forms: researchers may code variables, specify models, or exclude observations in ways that intentionally or unintentionally favor a particular conclusion. The file drawer problem is that entire studies go unreported simply because their results do not reach conventional thresholds for statistical significance. Rosenthal (1979) writes, "The extreme view of the 'file drawer problem' is that journals are filled with the 5% of the studies that show Type I errors, while the file drawers are filled with the 95% of the studies that show nonsignificant results." A growing body of evidence suggests that fishing and file drawer problems are widespread in the social sciences (Brodeur et al. 2016; Franco, Malhotra, and Simonovits 2014, 2016; Gerber and Malhotra 2008), undercutting the reliability of published research findings (Open Science Collaboration 2015).

Concerns about the fishing and file drawer problems have sparked a movement to promote transparency in social science research (Miguel et al. 2014; Nosek et al. 2015). One recent innovation is the public archiving of pre-analysis plans (PAPs) that specify details of the analysis (e.g., statistical methods, sample exclusions, outcome measures, covariates, and subgroup definitions) before the researchers see unblinded outcome data.[1] Deviations from the plans are not prohibited, but "when such deviations arise they [should] be highlighted and the effects on results reported" (Humphreys, Sanchez de la Sierra, and van der Windt 2013, 13).

In principle, PAPs have four main advantages. First, prespecification limits the extent to which researchers can make decisions that consciously or unconsciously tilt a study toward a desired result (Rubin 2007; Tukey 1993). Second, the validity of frequentist statistical inference (standard errors, confidence intervals, p-values, and significance tests) hinges on the assumption that the analysis follows a pre-specified strategy (Simmons, Nelson, and Simonsohn 2011; Tukey 1993). Third, publicly archived PAPs enable readers to see which analyses were pre-specified and to take that into account when assessing the credibility of results (Casey, Glennerster, and Miguel 2012; Chan et al. 2013; Freedman 2008, 2010; Humphreys, Sanchez de la Sierra, and van der Windt 2013; Monogan 2013, 2015; Tukey 1993). Fourth, PAPs are a prerequisite for a new model for publishing called Registered Reports (Chambers et al. 2014), in which "articles with prespecified analysis plans would be accepted in principle *before* the study was conducted" (Nyhan 2015).

In practice, researchers have found that PAPs have important benefits but can be challenging to write. On the one hand, writing a PAP can help researchers clarify their own thinking about research design and data collection, and getting sponsors or partners on board can protect against pressures for ex post changes in the analysis (Casey, Glennerster, and Miguel 2012; Humphreys, Sanchez de la Sierra, and van der Windt 2013; McKenzie 2012; Olken 2015). On the other hand, detailed PAPs are time-consuming to compose, and PAPs can easily fail to cover strokes of good or bad fortune, such as new data sources becoming available (Humphreys, Sanchez de la Sierra, and van der Windt 2013) or a school being hit by lightning (Olken 2015). As Humphreys, Sanchez de la Sierra, and van der Windt (2013, 11) write:

https://doi.org/10.1017/S1049096516000810 Published online by Cambridge University Press

Indeed, many things may go wrong that can lead to model changes in the analysis phase. *Ex ante* one may not know whether one will suffer from noncompliance, attrition, missing data, or other problems such as flaws in the implementation of randomization, flaws in the application of treatment, errors in data collection, or interruptions of data collection. Any of these possibly unanticipated features of the data could require fixes in the analysis stage. In each particular case, one could in principle describe precisely how to handle different data structures, but in the absence of an off-the-shelf set of best practices for all these issues, such efforts towards complete specification are likely to be onerous.

Keeping in mind Voltaire's aphorism that the best is the enemy of the good (O'Donoghue and Rabin 2001), we suggest developing off-the-shelf sets of *good* practices for *some* issues. A research group can specify a set of default practices, which we call a "standard operating procedure" (SOP), to guide decisions that have not been made explicit in a PAP. The SOP can support and flesh out PAPs, making them easier to write. The SOP should not replace PAPs, nor should it override the explicit decisions in PAPs. Rather, it can lighten the burden of preparing PAPs, especially when experimental opportunities arise suddenly and require researchers to make plans under tight deadlines.

This article describes some of the potential benefits of SOPs and offers an example from our research group (focusing on the analysis of randomized experiments) that others are welcome to adapt to suit their own needs.[2]

BENEFITS OF A STANDARD OPERATING PROCEDURE

Here are some scenarios where an SOP can provide guidance in the event that the PAP has not explicitly addressed the issue. Each scenario raises a question that a PAP could easily fail to anticipate.

- A project sponsor reveals to you that if a particular unit had not been assigned to treatment, the sponsor would have canceled the experiment. Thus, although treatment assignment was randomized, not every randomization would have yielded a reportable study. Should you still report the results, and if so, how should you analyze the data?
- After treatment has begun, you learn that some subjects were randomly assigned more than once. (For example, when applicants for a social program are randomly assigned as their applications are processed, randomization may go on for months or years, and in unusual cases, a persistent applicant who was originally assigned to the control group may later succeed in getting assigned to treatment.) How should their data be analyzed?
- You are conducting a randomized experiment to study the persuasive effects of a telephone canvassing effort, and have specified in the PAP that you will use an instrumental variables method (Angrist, Imbens, and Rubin 1996) to estimate the average effect of contact on those who were contacted. In the following situations, should the subject be coded as "contacted"?
  - The subject hung up right after the canvasser's initial greeting.
  - The canvasser never spoke to the subject but left a message with a housemate.
  - No one answered the phone, but the canvasser called from a number with a recognizable caller ID that identified the campaign.
- After playing behavioral games in a lab experiment, some subjects indicate that they had previously played several of the same games a few weeks earlier in a different lab. Should these subjects be excluded from the analysis?

An SOP can codify a research group's standing decisions on such issues, as well as others that are more routinely encountered, such as whether to report a one-tailed or a two-tailed test or how to handle missing covariate values. By specifying these decisions in the SOP, researchers eliminate the need to state them again and again when writing PAPs. Just as important, the SOP protects the researcher who might otherwise neglect to specify the procedure in a PAP. The SOP makes recurrent practices explicit and documents them ex ante so that researchers do not have to contend ex post that they were implicit.

DEVELOPING AND UPDATING AN SOP

Developing an SOP takes some up-front work, but we think that in the long run, the investment will be more helpful than onerous.

The SOP makes recurrent practices explicit and documents them ex ante so that researchers do not have to contend ex post that they were implicit.

To save time, one research group can borrow another group's existing SOP and modify it to fit their own needs and preferences. Different groups can collaborate on SOPs and learn from each other.

SOPs can be amended to reflect methodological innovations and lessons from experience. However, readers need some assurance that changes to SOPs are not just another form of data fishing. We suggest that each PAP either include the SOP as an appendix, or reference a specific SOP document that is archived and time-stamped in the same registry as the PAP. If an analysis follows the pre-registered PAP and SOP, it is clearly pre-specified. If it is guided by later amendments to the SOP, it falls into what Humphreys, Sanchez de la Sierra, and van der Windt (2013, 18) call "a gray zone in which analysis may stay true to the intent of the registered design but the defense of the details of implementation must be provided *ex post* rather than *ex ante*."[3] Pre-specified, gray-zone, and exploratory analyses can all be valuable, but readers need to know which is which.

Of course, any SOP will have gaps. When situations arise that are covered neither by the PAP nor by the SOP, we would still like to protect against the risk that research decisions will be influenced by their likely effects on results. One possible strategy is to consult a "jury" of colleagues who cannot see the unblinded outcome data or any information that might suggest whether a particular decision would make the estimated effects bigger or

https://doi.org/10.1017/S1049096516000810 Published online by Cambridge University Press

smaller. To make efficient use of jurors' time and expertise, such a jury might be invited to make binding decisions on a series of questions that were not anticipated by the PAP or SOP. The reasoning behind these decisions should be documented, and, if appropriate, the SOP should be amended to cover similar situations in the future. Jury decisions can be made public, so that the scholarly community can monitor and review the rulings, perhaps applying them to similar cases that may arise in other labs. In time, as experience and "common law" decisions accumulate, SOPs and the decision rules they embody will gradually become more comprehensive in scope.

We do not regard all of the defaults in our SOP as clearly superior to the alternatives. For example, in the section on covariate adjustment, we recommend that covariates be pre-specified "on the basis of their expected ability to help predict outcomes," give rules of thumb for the maximum number of covariates, and suggest how a jury can be used in exceptional cases (e.g., when a new source of baseline data becomes available after random assignment). We considered the alternative of adopting automated model selection methods, but would like to see more evidence that (1) valid confidence intervals can be constructed when such methods are used and (2) the benefits of such methods (possible improvements in precision) outweigh the costs (increased computing time, possible loss of transparency to non-expert readers). This is just one example of a topic where, as the literature evolves and experience accumulates, our SOP may evolve as well.

We believe that by building, borrowing, and discussing SOPs, researchers can share useful ideas about methodological issues and bolster the contributions of PAPs toward improving transparency.

OVERVIEW OF OUR CURRENT SOP

Our SOP (Lin, Green, and Coppock 2015) can viewed on GitHub, a web-based repository platform that allows us to publicly archive previous versions with tracked changes and allows users to post requests for additional issues to be addressed. The document can be downloaded without a GitHub account.

The principal motivation for the SOP is to support PAPs in pre-specifying analyses and credibly protecting against data fishing. Thus, the SOP focuses on data analysis, not experimental design, and it specifies our fallback plans for various analytic issues in case these were not addressed in the PAP. It is a document of default practices, not a manual of recommended practices. Our PAPs may deviate from the SOP when we believe a different approach is more appropriate for a particular study. Each PAP will include a statement that for any decisions not explicitly specified in the PAP, we plan to follow the SOP.

The SOP is a living document and will be expanded over time. Currently, it addresses several general topics (e.g., attrition, noncompliance, and use of covariates), some nonstandard situations we have encountered (e.g., learning that some subjects were randomly assigned more than once), and some issues specific to voter turnout experiments (e.g., how to code contact in canvassing experiments), survey experiments (e.g., whether to exclude subjects who are not paying attention to the content of questions), or lab experiments (e.g., whether to exclude subjects who indicate in a debriefing session that they discerned the purpose of the experiment). It does not attempt to cover all issues that may be important in analyzing experimental data.

For example, so far it does not address the multiple comparisons problem (Anderson 2008; Efron 2010; Westfall, Tobias, and Wolfinger 2011)—not because we think this issue is unimportant, but because we do not have an off-the-shelf recommendation for handling it. The multiple comparisons problem becomes more important as the number of outcome measures, treatments, or subgroups analyzed grows. Other researchers may find it useful to codify their multiple-comparisons practices in SOPs, especially if they typically analyze many outcome measures in a single study. We look forward to learning from their approaches and may address the issue in PAPs for specific projects and, if appropriate, a future version of our SOP.

Our SOP intentionally uses some arbitrary thresholds. For example, in several places in the sections on noncompliance and attrition, we specify statistical tests to compare baseline characteristics across treatment arms and write that "p-values below 0.05" will be considered evidence of noncomparability or asymmetric attrition, triggering changes in the analysis strategy. It may be wiser to pre-specify a rule based on substantive rather than statistical significance, and we may do so in PAPs, making use of subject-matter knowledge or simulations based on relevant data. However, the purpose of the SOP is to provide a fallback that constrains the analyst's discretion if the PAP does not address the issue, and a specific but arbitrary threshold serves this purpose more effectively than vague but judicious guidance.

In sharing our SOP, we are not seeking to persuade other research groups to adopt the same default practices we have chosen. In fact, we welcome debate and discussion about these practices and more opportunities to learn from other researchers' choices. We believe that by building, borrowing, and discussing SOPs, researchers can share useful ideas about methodological issues and bolster the contributions of PAPs toward improving transparency.

ACKNOWLEDGMENTS

Many thanks to Peter Aronow, Susanne Baltes, Jake Bowers, Alex Coppock, Al Fang, Macartan Humphreys, Hal Lin, Berk Özler, Anselm Rink, Cyrus Samii, Michael Schwam-Baird, Uri Simonsohn, Amber Spry, Dane Thorley, Anna Wilke, José Zubizarreta, and the editors and reviewers for helpful comments and discussions. ■

NOTES

1. One of us (Lin) worked at program evaluation firms that were already pre-specifying analyses of social experiments in the early 1990s. However, the public registration and archiving of time-stamped PAPs in the social sciences is a recent development. For valuable discussions, see Casey, Glennerster, and Miguel (2012), McKenzie (2012), Monogan (2015), Nyhan (2015), and the symposia in *Political Analysis* (Winter 2013) and *Journal of Economic Perspectives* (Summer 2015).

https://doi.org/10.1017/S1049096516000810 Published online by Cambridge University Press

2. Although we focus here on experiments, we think SOPs could also be useful for prospective observational studies. Monogan (2013) and Rubin (2007) discuss the benefits of pre-specification in observational studies.
3. A reviewer helpfully pointed us to the work of Bidwell, Casey, and Glennerster (2015), who provided "track changes" versions of their PAPs. The track-changes versions contain "a transparent record of every single change" made to the PAPs after the initial registration. We see this as a promising approach that could be used in combination with an SOP. One of the aims of an SOP is to reduce the need for changes to PAPs. If an analysis is guided by amendments to either the PAP or the SOP that were made after seeing the data, it falls into the "gray zone" described by Humphreys, Sanchez de la Sierra, and van der Windt (2013).

REFERENCES

Anderson, Michael L. 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association* 103 (484): 1481–1495.

Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91 (434): 444–455.

Bidwell, Kelly, Katherine Casey, and Rachel Glennerster. 2015. "The Impact of Voter Knowledge Initiatives in Sierra Leone." AEA RCT Registry. https://www.socialscienceregistry.org/trials/26.

Brodeur, Abel, Mathias Lé, Marc Sangnier, and Yanos Zylberberg. 2016. "Star Wars: The Empirics Strike Back." *American Economic Journal: Applied Economics* 8 (1): 1–32.

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.

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.

Chan, An-Wen, Jennifer M. Tetzlaff, Peter C. Gøtzsche, Douglas G. Altman, Howard Mann, Jesse A. Berlin, Kay Dickersin, Asbjørn Hróbjartsson, Kenneth F. Schulz, Wendy R. Parulekar, Karmela Krleža-Jeric, Andreas Laupacis, and David Moher. 2013. "SPIRIT 2013 Explanation and Elaboration: Guidance for Protocols of Clinical Trials." *BMJ* 346: e7586.

Efron, Bradley. 2010. *Large-Scale Inference: Empirical Bayes Methods for Estimation, Testing, and Prediction*. New York: Cambridge University Press.

Franco, Annie, Neil Malhotra, and Gabor Simonovits. 2014. "Publication Bias in the Social Sciences: Unlocking the File Drawer." *Science* 345 (6203): 1502–1505.

———. 2016. "Underreporting in Psychology Experiments: Evidence from a Study Registry." *Social Psychological and Personality Science* 7 (1): 8–12.

Freedman, David A. 2008. "Oasis or Mirage?" *Chance* 21 (1): 59–61.

———. 2010. "Survival Analysis: An Epidemiological Hazard?" In *Statistical Models and Causal Inference: A Dialogue with the Social Sciences*, ed. David Collier, Jasjeet S. Sekhon, and Philip B. Stark, 169–192. New York: Cambridge University Press.

Gerber, Alan and Neil Malhotra. 2008. "Do Statistical Reporting Standards Affect What Is Published? Publication Bias in Two Leading Political Science Journals." *Quarterly Journal of Political Science* 3 (3): 313– 326.

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.

Lin, Winston, Donald P. Green, and Alexander Coppock. 2015. "Standard Operating Procedures for Don Green's Lab at Columbia." https://github.com/acoppock/Green-Lab-SOP.

McKenzie, David. 2012. "A Pre-Analysis Plan Checklist." World Bank. http://blogs.worldbank.org/impactevaluations/a-pre-analysis-plan-checklist.

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.

Monogan, James E., III. 2013. "A Case for Registering Studies of Political Outcomes: An Application in the 2010 House Elections." *Political Analysis* 21 (1): 21–37.

———. 2015. "Research Preregistration in Political Science: The Case, Counterarguments, and a Response to Critiques." *PS: Political Science and Politics* 48 (03): 425–429.

Nosek, B. A., G. Alter, G. C. Banks, D. Borsboom, S. D. Bowman, S. J. Breckler, S. Buck, C. D. Chambers, G. Chin, G. Christensen, M. Contestabile, A. Dafoe, E. Eich, J. Freese, R. Glennerster, D. Goroff, D. P. Green, B. Hesse, M. Humphreys, J. Ishiyama, D. Karlan, A. Kraut, A. Lupia, P. Mabry, T. Madon, N. Malhotra, E. Mayo-Wilson, M. McNutt, E. Miguel, E. L. Paluck, U. Simonsohn, C. Soderberg, B. A. Spellman, J. Turitto, G. VandenBos, S. Vazire, E. J. Wagenmakers, R. Wilson, and T. Yarkoni. 2015. "Promoting an Open Research Culture." *Science* 348 (6242): 1422–1425.

Nyhan, Brendan. 2015. "Increasing the Credibility of Political Science Research: A Proposal for Journal Reforms." *PS: Political Science and Politics* 48 (S1): 78–83.

O'Donoghue, Ted and Matthew Rabin. 2001. "Choice and Procrastination." *Quarterly Journal of Economics* 116 (1): 121–160.

Olken, Benjamin A. 2015. "Promises and Perils of Pre-Analysis Plans." *Journal of Economic Perspectives* 29 (3): 61–80.

Open Science Collaboration. 2015. "Estimating the Reproducibility of Psychological Science." *Science* 349: aac4716.

Rosenthal, Robert. 1979. "The 'File Drawer Problem' and Tolerance for Null Results." *Psychological Bulletin* 86 (3): 638–641.

Rubin, Donald B. 2007. "The Design *versus* the Analysis of Observational Studies for Causal Effects: Parallels with the Design of Randomized Trials." *Statistics in Medicine* 26 (1): 20–36.

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.

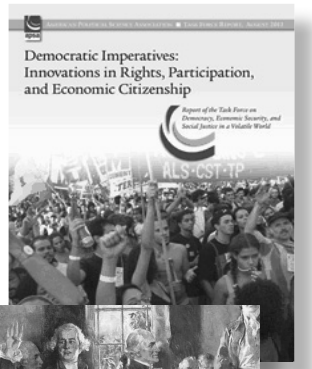
Tukey, J. W. 1993. "Tightening the Clinical Trial." *Controlled Clinical Trials* 14 (4): 266–285.

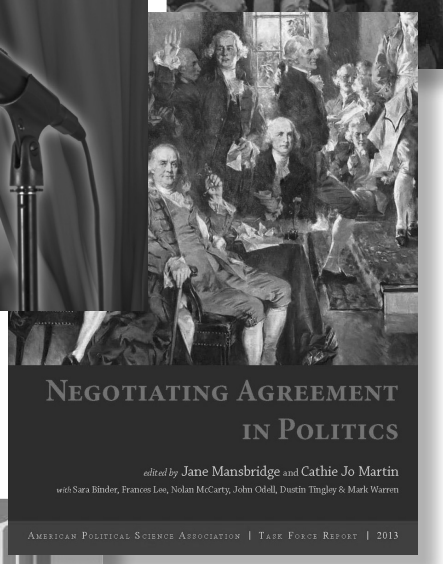
Westfall, Peter H., Randall D. Tobias, and Russell D. Wolfinger. 2011. *Multiple Comparisons and Multiple Tests Using SAS*. 2nd ed. Cary, NC: SAS Institute.

https://doi.org/10.1017/S1049096516000810 Published online by Cambridge University Press

APSA Task Force Reports



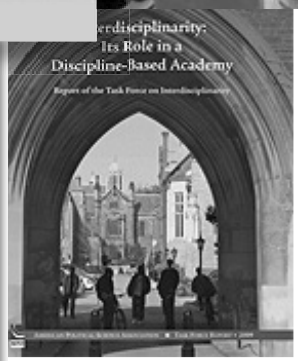












The APSA task force reports seek to expand the public presence of political science by:

- applying political science research to public policy issues,
- and sharing expertise on the most significant topics of public concern.

APSA's current task force report!

Let's Be Heard! How to Better Communicate Political Science's Public Value,
in a special issue of PS

Read the complete APSA task force report, which identifies how individual scholars and professional organizations can make political science's insights and discoveries more accessible, more relevant, and more valuable to more people.

These reports and others are also available as FREE .pdf downloads!

Visit **www.apsanet.org/reports** to obtain your APSA task force publications.

apsa AMERICAN POLITICAL SCIENCE ASSOCIATION

For questions or concerns, please contact publications@apsanet.org.

https://doi.org/10.1017/S1049096516000810 Published online by Cambridge University Press