Openness in Political Science: Data Access and Research Transparency

Introduction

Arthur Lupia, *University of Michigan*
Colin Elman, *Syracuse University*

In 2012, the American Political Science Association (APSA) Council adopted new policies guiding data access and research transparency in political science. The policies appear as a revision to APSA's *Guide to Professional Ethics in Political Science.* The revisions were the product of an extended and broad consultation with a variety of APSA committees and the association's membership.[1]

After adding these changes to the ethics guide, APSA asked an Ad Hoc Committee of scholars actively discussing data access and research transparency (DA-RT) to provide guidance for instantiating these general principles in different research traditions. Although the changes in the ethics guide articulate a single set of general principles that apply across the research traditions, it was understood that different research communities would apply the principles in different ways. Accordingly, the DA-RT Ad Hoc Committee formed sub-committees to draft more fine-grained guidelines for scholars, journal editors, and program managers at funding agencies who work with one or more of these communities. The subcommittees have produced circulation drafts for APSA members' review and comment. The drafts are titled *Guidelines for Data Access and Research Transparency in the Quantitative Tradition* and *Guidelines for Data Access and Research Transparency in the Qualitative Tradition*[2] and are attached as Symposium Appendices A and B.

This article is the lead entry of a *PS: Political Science and Politics* symposium on the ethics guide changes described above, the continuing DA-RT project, and what these endeavors mean for individual political scientists and the discipline. Its content is as follows. In the first section, we offer a brief history of how the ethics guide changes came about and our understanding of the motivations of the diverse group of scholars who work on the DA-RT initiative. In the second section,

we present the changes to the ethics guide. In the third section, we work from these changes to offer a broader argument about the value of greater openness to individual political scientists and to the discipline. We conclude by providing a brief summary of themes developed in the symposium's seven subsequent articles and inviting feedback.

With this content in mind, we want to draw your attention to the fact that DA-RT is an open endeavor. While we are listed as authors on this particular article, the progress made in this domain in recent years is the result of the effort of numerous social scientists. In addition to being open, DA-RT is an ongoing effort in which any political scientist can participate. We hope that you will find in this symposium ways to increase the value and impact of your efforts as teachers, researchers, and public servants.

HISTORY

Political science is a diverse discipline comprising multiple, and sometimes seemingly irretrievably insular, research communities. We could spend much of this introduction (indeed fill several issues of the journal) on the sociology of academic disciplines and why they tend to fragment. But recent discussions about openness are a rare and welcome example of dissimilar scholars finding opportunities for collaboration and common action.

Several years ago, APSA's governing council, under the leadership of president Henry E. Brady, began an examination of research transparency. Its initial concerns were focused on the growing concern that scholars could not replicate a significant number of empirical claims that were being made in the discipline's leading journals. There were multiple instances where scholars would not, or could not, provide information about how they had selected cases, or how they had derived a particular conclusion from a specific set of data or observations. Other scholars refused to share data from which others could learn. Still other scholars would have been willing to share their data, but failed to archive them in effective ways, making the information unavailable for subsequent inquiries.

As political scientists described such episodes to each other, they realized that scholars from different methodological and substantive subfields were having similar experiences and

conversations. In a wide range of circumstances, professional customs and incentives for sharing information and data were less well developed than those for producing knowledge claims. An unusually diverse set of political scientists identified common concerns and aspirations, both in their reasons for wanting greater openness and in the benefits that new practices could bring.

What is political scientists' shared interest in openness? As Elman and Kapiszewski (2014) note, openness is best understood as a meta standard that applies to all social inquiry. All rule-based social inquiry is based on three notions: first, scholarly communities hold shared and stable beliefs that research designed and conducted in particular ways possesses certain characteristics. Second, both the conduct of social inquiry and the written products that represent its conclusions are designed to capture those characteristics. Finally, for any given piece of research in a particular tradition, the ability of scholars to claim the underlying warrants depends on their showing that it was designed and conducted in accordance with those rules. The view that social science is a group activity, requiring inter-subjective knowledge being created using public processes that are warranted to add value, is common to virtually every scholarly tradition.

ars who produce knowledge claims want others to have a rationale for believing those claims. Therefore, DA-RT operates from a "community standards" approach, where optimal means of data sharing and research transparency respect and build from the challenges and opportunities that characterize various research traditions. Because social scientists use different methods, how a knowledge claim achieves credibility and legitimacy depends on the type of work. For all research traditions in political science, our main focus is to better equip its scholars with incentives and mechanisms for making their knowledge claims easier for others to interpret and assess accurately.

That said, the shared commitment to openness places limits on practices that DA-RT can endorse. For example, DA-RT rules out claims about the credibility and legitimacy of scientific claims based solely on personality cults or on raw exercises in power (i.e., "the claim is true because my minions and I so testify"). What distinguishes scientific claims from others is the extent to which scholars attach to their claims publicly available information about the steps that they took to convert information from the past into conclusions about the past, present, or future.

The credibility of scientific claims comes, in part, from the fact that their meaning is, at a minimum, available for other

> *Therefore, DA-RT operates from a "community standards" approach, where optimal means of data sharing and research transparency respect and build from the challenges and opportunities that characterize various research traditions.*

Communities have very different beliefs about what constitutes useful knowledge and how such value is to be obtained. That said, there is substantial overlap about which attributes of openness contribute to accurate inter-subjective knowledge transfer. Our prescriptive methodologies all involve extracting information from the social world, analyzing the resulting data, and reaching a conclusion based on a combination of the evidence and its analysis. No matter whether the research is, for example, ethnographic field work, a laboratory experiment, or the statistical analysis of a large data set, they all combine assumptions, decisions, and actions that produce evidence and analysis. Sharing information about these assumptions, decisions, and actions is necessary for scholars to place one another's meanings in a legitimizing context. DA-RT is motivated by this premise—the principle that sharing data and information fuels a culture of openness that promotes effective knowledge transfer.

This justification for openness (the desire to establish a knowledge claim's validity) and its general content (showing both evidence and analysis) are epistemically neutral. They apply wherever scholars seek to use a shared logic of inquiry to reach evidence-based conclusions. To this end, a critical attribute of DA-RT is that it does not impose a uniform set of standards on political scientists. Instead, it begins from a simple premise about credibility and legitimacy. In short, schol-

scholars to rigorously evaluate. In other words, the reason to believe a scientist's claim is not because he or she wears a lab coat, have a PhD, or have published a widely viewed paper in the past. Appeals to personality or faith, which facilitate information transmission in other domains, are not supposed to be required to access the content of a scientific claim. A claim's perceived legitimacy is grounded in the fact that the results are the product of publicly described processes that in turn are based on a stable and shared set of beliefs about how knowledge is produced. Such open access to the origins of others' claims is the hallmark of scientific ways of knowing.

Accordingly, when social scientists fail to document their assumptions, decisions, and actions and are unwilling or unable to share this information with others, it limits others' abilities to understand the meaning of the scientists' claims. When such failures are frequent in a research community, the credibility and legitimacy of the community as a whole are imperiled. Across the sciences, questions about data sharing and research transparency are now being increasingly and vigorously addressed. Advances in electronic communication not only expose scholars to a wider set of knowledge claims, but also give them reasons to expect that data and inferential information can be made more readily available. DA-RT is one of several efforts in the social sciences to advance the cause of transparency.

DA-RT's distinction is that it is focused on political science. Our goal is to provide, through a community standards approach, individual scholars of every epistemic tradition opportunities for greater openness, transparency, legitimacy, and credibility. This goal has motivated a diverse set of scholars to contribute to the DA-RT project. These scholars have developed a wide range of mechanisms to increase professional incentives for data sharing and research transparency. They have also worked to make such activities easier for a growing range of scholars. DA-RT is a movement that anyone interested in political science can join.

The credibility of scientific claims comes, in part, from the fact that their meaning is, at a minimum, available for other scholars to rigorously evaluate. In other words, the reason to believe a scientist's claim is not because they wear a lab coat, have a PhD, or have published a widely viewed paper in the past. Appeals to personality or faith, which facilitate information transmission in other domains, are not supposed to be required to access the content of a scientific claim.

ETHICS GUIDE CHANGES

APSA's ethics guidelines now state that "researchers have an ethical obligation to facilitate the evaluation of their evidence-based knowledge claims through data access, production transparency, and analytic transparency so that their work can be tested or replicated." The three constitutive elements are defined as follows:

6.1 *Data access*: Researchers making evidence-based knowledge claims should reference the data they used to make those claims. If these are data they themselves generated or collected, researchers should provide access to those data or explain why they cannot.

6.2 *Production transparency*: Researchers providing access to data they themselves generated or collected, should offer a full account of the procedures used to collect or generate the data.

6.3 *Analytic Transparency*: Researchers making evidence-based knowledge claims should provide a full account of how they draw their analytic conclusions from the data, i.e., clearly explicate the links connecting data to conclusions.

6.4 Scholars may be exempted from Data Access and Production Transparency in order to (A) address well-founded privacy and confidentiality concerns, including abiding by relevant human subjects regulation; and/or (B) comply with relevant and applicable laws, including copyright. Decisions to withhold data and a full account of the procedures used to collect or generate them should be made in good faith and on reasonable grounds. Researchers must, however, exercise appropriate restraint in making claims as to the confidential nature of their sources, and resolve all reasonable doubts in favor of full disclosure.

6.5 Dependent upon how and where data are stored, access may involve additional costs to the requesting researcher.

6.6 Researchers who collect or generate data have the right to use those data first. Hence, scholars may postpone data access and production transparency for one year after publication of evidence-based knowledge claims relying on those data, or such period as may be specified by (1) the journal or press publishing the claims, or (2) the funding agency supporting the research through which the data were generated or collected.

6.7 Nothing in this section shall require researchers to transfer ownership or other proprietary rights they may have.

6.8 As citizens, researchers have an obligation to cooperate with grand juries, other law enforcement agencies, and institutional officials. Conversely, researchers also have a professional duty not to divulge the identity of confidential sources of information or data developed in the course of research, whether to governmental or nongovernmental officials or bodies, even though in the present state of American law they run the risk of suffering an applicable penalty.

6.9 Where evidence-based knowledge claims are challenged, those challenges are to be specific rather than generalized or vague. Challengers are themselves in the status of authors in connection with the statements that they make, and therefore bear the same responsibilities regarding data access, production transparency, and analytic transparency as other authors.

While data access and research transparency are the "default" settings in the new guidelines, these expectations are contingent on the author not putting people at risk or breaking the law. Hence concerns about human subjects protections and copyright limitations are accounted for in the new language.

With these changes, APSA's ethics guide is more consistent with current and emerging standards across the sciences. Where APSA's previous language emphasized making data accessible *only* when findings were challenged, the new guidelines recognize data access and research transparency as an indispensable part of the research endeavor. It is also critical to notice that the updated language is epistemically neutral: it respects the integrity of different research traditions, and the diverse data collection and analytic steps that they take.

HOW POLITICAL SCIENCE BENEFITS FROM INCREASED OPENNESS

A more rigorous and self-conscious approach to openness promises several benefits to political scientists. One way to categorize these benefits is with respect to the different audiences for political science scholarship.

First, and most obviously, transparency offers an opportunity for members of a particular research community to understand and assess their own scholarship. Data sharing and research transparency allow a researcher's audience to evaluate claims and form an evidentiary and logical basis for treating the claims as valid.

The most widespread (although as we note below, not universal) way that this principle is pursued is through replication. For subfields that hold that inferential procedures are

puter simulations of war. Higher standards of data access and research transparency will make cross-border understanding more attainable.

Other audiences are not focally involved in research. Instead, they want to use research claims as the basis of action. Teachers, for example, want to use the claims for pedagogical purposes. Whether demonstrating substantive arguments about aspects of the social world, or training students to use research techniques, teaching is substantially improved by the availability of exemplary scholarship, with its data and reasoning on display.

Public and private sector decision makers comprise another audience. Their main interest is in using knowledge claims to improve the effectiveness and efficiency of valuable endeavors. Greater openness gives such audiences increased oppor-

> *For subfields that hold that inferential procedures are repeatable, openness is a necessary condition for replication. For these communities, replication of another's claims provides increased confidence in the validity of that work.*

repeatable, openness is a necessary condition for replication. For these communities, replication of another's claims provides increased confidence in the validity of that work. When subfields have such confidence, they can devote their attention to evaluating competing theories of important phenomena. If, by contrast, opportunities for replication are diminished because of poor data availability or incomplete accounts of how results were reached, it is impossible to determine the strength or robustness of findings—which makes confidence harder to build.

Members of other research communities do not validate one another's claims by repeating the analyses that produced them. In these communities, the justification for transparency is not replication, but understandability and persuasiveness. The more material scholars make available, the more that they can accurately relate such claims to a legitimating context. When readers are empowered to make sense of others' arguments in these ways, the more pathways exist for readers to believe and value knowledge claims. Whether scholars privilege replication, context-specificity, or other ways of evaluating the meaning of a knowledge claim, sharing information that allow such evaluations facilitates knowledge transfer. Hence, research openness is a broader ideal, and one from which scholars can benefit regardless of which viewpoint they take on replication.

Second, openness is beneficial for scholars outside the immediate community in which the research is located. Political science is a methodologically diverse discipline, and we are sometimes unable to appreciate how other social scientists generate their conclusions. Mathematical modelers, for example, often know very little about how cases are selected in participant observation studies—and many people who seek meaning in texts have a limited understanding of how other social scientists try to seek meaning from surveys or com-

tunities to understand how the claims relate to their aspirations. As Lupia (2014) notes, many decision makers value information whose veracity they can readily defend in politicized contexts. These decision makers find claims whose origins are available and accessible more valuable informational currency than claims whose foundations are hidden.

Beyond general openness, data sharing provides an important additional benefit—it allows secondary analysis. Shared data can be a valuable public good. Secondary analysts can use data in ways that data originators did not. In the best-case scenario, secondary data analyses allow authors to derive meaning from data that need not have occurred to the original researcher. When scholars can use research materials in these diverse ways, the data can become more valuable to science and society. Instead of a dataset producing one set of insights, data sharing gives other scholars the ability to multiply datas' value.

Many of these benefits of openness are widely known. We have found, however, that while the goals of greater data sharing and research transparency are generally accepted, they are less often followed in practice. Most political scientists to whom we have spoken find nothing radical or challenging about the notion that they show the information and analysis underpinning their evidence-based claims. But as the articles in this symposium show, there are multiple instances in which individual actions do not live up to our shared aspirations.

One challenge is that quantitative and qualitative research traditions lack clearly specified guidelines as to what kinds of data and research information should be shared. Compounding this problem is a lack of professional incentives for documenting the evidentiary and logical foundations of knowledge claims, the temporal and monetary expense that can be involved in archiving research materials, and the potential for embarrassment that can come from having one's work

reexamined. These are all substantial headwinds confronting transparency movements. The question for individual investigators and the discipline as a whole is whether we can derive the benefits of greater openness while recognizing, and then minimizing, the costs.

The contributions to this symposium are motivated principally by such challenges and questions.

TOPIC OF THE SYMPOSIUM: NEXT STEPS IN DATA ACCESS AND RESEARCH TRANSPARENCY

This symposium contains seven articles on DA-RT-related activities. Each article is written by scholars interested in investigating the benefits of greater openness and offering ideas about how to make data access and research transparency more viable and incentive-compatible activities for all political scientists. The distinct contribution of each article to this cause is to identify where potential gains from openness are appar-

sions that we make about graduate student training. Isihiyama, lead editor of the *American Political Science Review*, describes the different ways in which journals are adapting to calls for greater openness. He concludes by offering a number of different ways that journals can better address demands for greater openness, including replication studies.

In many areas of the discipline, there are limited incentives to increase openness. At the same time, there are multiple levers the discipline can pull to increase openness's incentive compatibility for the purpose of augmenting political science's legitimacy. These levers include changing disciplinary norms so that data production is valued for promotion and tenure, developing software tools to lower barriers to entry for curating data (for example, the Active Citation Editor (ACE) and the Live Active Citation Editor (LACE) in qualitative research), and incentivizing graduate students for greater openness from the beginning of their careers.[3]

The distinct contribution of each article is to identify where potential gains from openness are apparent but not yet fully realized. In each case, the authors seek to reconcile individual incentives, existing norms, and possible ways of changing rewards and technology to increase the frequency and effect of greater openness.

ent but not yet fully realized. In each case, the authors seek to reconcile individual incentives, existing norms, and possible ways of changing rewards and technology to increase the frequency and effect of greater openness.

This introduction is followed by two articles focused on qualitative research. Colin Elman and Diana Kapiszewski discuss how openness is instantiated differently in diverse qualitative research traditions. They illustrate this discussion with a brief account of some concerns that arise when making process tracing research transparent. Andrew Moravscik shows how a practice called *active citation* can be implemented to increase the credibility and legitimacy of a wide range of qualitative research.

The next two articles (Arthur Lupia and George Alter, and Allan Dafoe) concentrate on large-N observational studies. Lupia and Alter discuss general opportunities for, and challenges to, increased openness that face quantitative scholars. Dafoe cites the benefits of sharing complete replication files for scholars who base conclusions on various forms of high-N statistical inference.

Rose McDermott focuses on experimental research. She discusses several innovative openness proposals in that domain including experimental registries—a system where scholars commit to publicizing their research designs before collecting data so that readers can better evaluate the meaning and generalizability of experimental results. The symposium concludes with articles by Thomas M. Carsey and John Ishiyama on the topic of how to implement critical elements of the DA-RT agenda. Carsey, director of the Odum Institute, describes new and emerging archiving opportunities and makes a strong argument for how the success of such opportunities is tied to deci-

Each contributor to this symposium offers creative ideas about how to move forward and each of their views has informed our own. Taken together, the articles make the case that openness is an indispensable element of credible research and rigorous analysis, and hence essential to both making and demonstrating scientific progress. These articles represent the great energy for increased credibility that a deeper and more sustained commitment to DA-RT principles can bring.

If you are not yet familiar with DA-RT, the changes to the ethics guide, and their implications for future activity in our discipline, then this symposium is a good place to learn more about these topics. Having engaged the materials, we hope that you will join our effort. Admission is free, and we can use all the help that you can offer. ∎

NOTES

1. The first DA-RT text was drafted by an Ad Hoc Committee, which consisted of Arthur Lupia (University of Michigan), Colin Elman (Syracuse University), George C. Alter (University of Michigan), Brian D. Humes (National Science Foundation), Diana Kapiszewski (Georgetown University), Rose McDermott (Brown University), Ron Rogowski (University of California, Los Angeles), S. Laurel Weldon (Purdue University), and Rick Wilson (Rice University). The suggested changes were reviewed and amended by APSA's Committee on Professional Ethics, Rights, and Freedoms, which consisted of Richard G.C. Johnston (University of British Columbia), Michael Lienesch (University of North Carolina, Chapel Hill), Marion Smiley (Brandeis University), Philip A. Schrodt (Pennsylvania State University), Sarah Birch (University of Essex), and Christian Davenport (University of Notre Dame). At the spring 2012 APSA Council Meeting in Chicago, the council adopted the language put forward by the Ethics Committee as APSA policy. The language was posted to APSANET and circulated to the membership. Following that consultation, the council at its October 2012 meeting formally voted to include the new language in the association's *Guide to Professional Ethics*.

2. To ensure continuity, and so that the process could benefit from the Ad Hoc Committee's expertise, the follow-on committees include a

combination of the original members and newly invited participants. The qualitative committee is chaired by Colin Elman (Syracuse University), and includes Diana Kapiszewski (Georgetown University), Rose McDermott (Brown University), Andrew Moravcsik (Princeton University), Brian Humes (National Science Foundation), Elizabeth Saunders (George Washington University), and Marc Trachtenberg (University of California, Los Angeles). The quantitative committee is chaired by George Alter (University of Michigan and Director of ICPSR), and includes Arthur Lupia (University of Michigan), Brian Humes (National Science Foundation), Gary King (Harvard University), Christopher Zorn (Pennsylvania State University), Rick K. Wilson (Rice University), Michael Alvarez (California Institute of Technology), Dara Strolovitch (University of Minnesota), Tom Carsey (University of North Carolina, Chapel Hill), and Valerie Martinez-Ebers (*APSR* and University of North Texas).

3. ACE and LACE are currently in development at Syracuse University's Qualitative Data Repository.

REFERENCES

Elman, Colin, and Diana Kapiszewski. 2014. "Data Access and Research Transparency in the Qualitative Tradition." *PS: Political Science and Politics* 47 (1): this issue.

Lupia, Arthur. 2014. "What Is the Value of Social Science? Challenges for Researchers and Government Funders." *PS: Political Science and Politics* 47 (1): this issue.

SYMPOSIUM AUTHORS

George Alter *is professor of history at the University of Michigan and director of the Inter-university Consortium for Political and Social Research. His research interests lie in the history of the family, demography, and economic history. Recent work explores demographic responses to economic hardship in Europe and East Asia, and the effects of childhood experiences on health in old age. He can be reached at altergc@umich.edu.*

Thomas M. Carsey *is the Pearsall Distinguished Professor of Political Science at the University of North Carolina at Chapel Hill. His research focuses on representation in US state and national politics, campaigns and elections, party polarization, and quantitative research methods. He also serves as Director of the Odum Institute for Research in Social Science, which operates a large social science data archive, and is the editor for the academic journal* State Politics and Policy Quarterly. *He can be reached at carsey@unc.edu.*

Allan Dafoe *is assistant professor of political science at Yale University. His research examines the causes of war, with emphases on the character and causes of the liberal peace, reputational phenomena such as honor and tests of resolve, and escalation dynamics. He can be reached at allan.dafoe@yale.edu.*

Colin Elman *is associate professor of political science, at the Maxwell School of Citizenship and Public Affairs, Syracuse University. His areas of interests are international relations, national security and qualitative methods. He is co-founder and director of the Institute for Qualitative and Multi-Method Research,* which offers intensive social science methods training, and co-editor of Cambridge University Press' *Strategies for Social Inquiry series. He is co-director of the Qualitative Data Repository. He can be reached at celman@maxwell.syr.edu.*

John Ishiyama, *University Distinguished Research Professor of Political Science at the University of North Texas and lead editor of* American Political Science Review, *is a comparative politics scholar, who specializes in political parties and democratization in post-communist Russian, East Central European, and African (particularly Ethiopian) politics. He has also done considerable work on ethnic conflict and politics (particularly the role played by ethnic parties) and on the scholarship of teaching and learning. He can be reached at john.ishiyama@unt.edu.*

Diana Kapiszewski *is assistant professor of government at Georgetown University. Her research focuses on comparative judicial politics and qualitative methods in political science. She is a co-author of* Field Research in Political Science *which will be published by Cambridge University Press in 2014. Kapiszewski was recently awarded the David Collier Mid-Career Achievement Award by the APSA Section for Qualitative and Multi-Method Research. She is co-director of the Qualitative Data Repository. She can be reached at dk784@georgetown.edu.*

Arthur Lupia *is the Hal R. Varian Collegiate Professor of Political Science at the University of Michigan. He has served on APSA's Governing Council, Executive Council and as its treasurer. He is president of the Midwest Political Science Association and chair of* American Association for the Advancement of Science's Social, Behavioral, and Economics Division. *He is principal investigator of EITM and has served as principal investigator of the American National Election Studies and TESS. He can be reached at lupia@umich.edu.*

Rose McDermott *is professor of political science at Brown University. Her main area of research revolves around political psychology in international relations. She has authored three books, co-edited two additional books, and has written numerous articles and book chapters on experimentation, evolutionary and neuroscientific models of political science, political behavior genetics and the impact of emotion on decision making. She can be reached at Rose_McDermott@brown.edu.*

Andrew Moravcsik *is professor of politics and director, European Union Program, at the Woodrow Wilson School at Princeton University. He has authored over 125 scholarly publications, including four books, on European integration, international relations theory, qualitative/historical methods, and other topics. He has served as trade negotiator for the US government, special assistant to the Deputy Prime Minister of the Republic of Korea, press assistant for the European Commission, editor of a Washington foreign policy journal, and on various policy commissions.at Princeton University. He can be reached at amoravcs@princeton.edu.*

Guidelines for Data Access and Research Transparency for Qualitative Research in Political Science[1]

In October 2012, the American Political Science Association (APSA) adopted new policies requiring transparency in political science research. The new policies have been integrated into Section 6 of the Association's *Guide to Professional Ethics, Rights and Freedoms* (and are reproduced in the Appendix to this document).

The new standards require researchers making evidence-based knowledge claims in their published work to provide data access, and engage in production transparency and analytic transparency.

- *Data access* requires authors to reference the data on which their descriptive and causal inferences and interpretations are based and, if they generated or collected those data, to make them available or explain why they cannot.

- *Production transparency* requires authors who collected and/or generated the data serving as evidence for their claims to explain the genesis of those data. Production transparency is necessary for other scholars to understand and interpret the data which authors have made available.

- *Analytic transparency* requires that authors demonstrate how they used cited data to arrive at evidence-based claims.

The promulgation of an APSA standard underscores a growing disciplinary (and multi-disciplinary) consensus that data access, production transparency and analytic transparency are all critical aspects of the research process. Transparency contributes to the credibility and legitimacy of political science research and facilitates the accumulation of knowledge. Assessing, critiquing, and debating evidence-based claims made in published research require access to the data cited to support them, documentation and metadata describing how those data were generated or collected, and an explanation of how the evidence and claims are connected. Providing access to data, and to documentation describing data generation or collection, also makes data more useful for testing new theories, for the development of new datasets and bodies of evidence, and for other forms of secondary data analysis.

Data access, production transparency, and analytic transparency are interconnected. Data access is a precondition for evaluating how data are used. Production transparency is a key prerequisite for evaluating author-provided data, and the connections that authors posit between those data and their inferences and interpretations. Conversely, one can more effectively evaluate an author's data generation or collection techniques (revealed through production transparency) when one knows for what analytical use the data are intended.

---

[1] We gratefully acknowledge the assistance of Louise Corti, Associate Director and Functional Director, UK Data Service, for helpful comments on earlier versions of this document.

This document is a resource for scholars, journal editors and academic evaluators (reviewers, funders or award committees) who seek assistance in satisfying these new data access and research transparency obligations in the context of qualitative research.[2] Accordingly, the document provides prospective guidance for meeting the obligations, as well as for retrospectively assessing whether they have been satisfied. While the new standards encourage as much data sharing and research transparency as possible, they should not be viewed in all-or-nothing terms: these activities often face friction, for example in the form of human subjects or copyright concerns. Sharing some data and being as transparent as possible, within those or other limits, will generally be better than doing neither at all.

The document's contents apply to all qualitative analytic techniques employed to support evidence-based claims, as well as all qualitative source materials.[3] No matter which qualitative techniques scholars use, research-tradition specific standards of transparency allow scholars to demonstrate the richness and rigor of qualitative work, and make clear its considerable contributions to knowledge accumulation and theory generation.

The Argument for Research Tradition-Specific Transparency Practices

The need for transparency in qualitative political science research derives from the fundamental principles which underlie social science as a community-based activity. Enhancing transparency both augments the quality of qualitative political science and increases its salience in and contributions to the discipline. Transparency is best achieved in qualitative political science in ways that preserve and honor that research tradition. We argue each of these points in turn.

Why Adopt Transparency Practices?

Transparency is an indispensable element of rule-bound intersubjective knowledge. Scholarly communities in the social sciences, natural sciences and evidence-based humanities can only exist if their members openly share evidence, results and arguments. Transparency allows those communities to recognize when research has been conducted rigorously, to distinguish between valid and invalid propositions, to

---

[2] A parallel set of guidelines intended primarily for quantitative data has also been developed. Of course, the guidance concerning ways in which research can be made more transparent offered in these documents is not exhaustive. In particular, nothing here is intended to prevent or discourage the development of more fine-grained requirements attuned to a particular subset of research, such as registering research designs involving experiments prior to conducting research with the aim of preventing publication and reporting bias.

[3] Such materials encompass traditional sources, such as primary textual documents and published primary sources; data from interviews, focus groups, or oral histories (in either audio or video form or transcripts from or summaries thereof); field notes (for instance from participant observation or ethnography); diaries and other personal records; and press clippings. The guidelines also apply to less conventional sources such as samples from bodies of secondary work; photographs; maps, posters and other representational work; and artwork.

better comprehend the subjective social understandings underlying different interpretations, to expand the number of participants in disciplinary conversations, and to achieve scientific progress.

To date, this fundamental attribute of community-based knowledge generation has played out in political science primarily in the realm of replicating quantitative research. In contrast to the situation in legal academia, historical studies, classical philology and some other disciplines, in qualitative political science transparency norms have been weak or non-existent. To be sure, citations and references in qualitative research appear to assure openness. Nevertheless, imprecision in citation, the high transaction costs of actually locating cited evidence, and the opacity of links between data and conclusions, combine to make the critical evaluation of descriptive and causal inferences or cumulative deepening of data analysis a rare event.

The aim of transparency is to make the rigor and power of good qualitative research more visible, allowing and empowering each consumer to identify such research, and facilitating the awarding of appropriate credit. Further, increasing the ease with which a larger number of scholars can critically engage with qualitative research, and the depth with which they can do so, makes it more likely that such work will be incorporated into scholarly discussion and debate, and future research. In all these ways, enhancing understanding of the processes and products of qualitative research facilitates the accumulation of knowledge.

Why an Approach to Transparency that is Specific to Qualitative Research?

Transparency in any research tradition – whether quantitative or qualitative – requires that scholars show they followed the rules of data collection and analysis that guide the specific type of research in which they are engaged. That conformity is foundational to the validity of the resulting interpretations and inferences and its demonstration is a key component of social science.

A shared commitment to openness, however, does not oblige all research traditions to adopt the same approach. Rather, transparency should be pursued in ways and for reasons that are consistent with the epistemology of the social inquiry being carried out. There are several reasons why qualitative scholars should not (and sometimes simply could not) adopt the transparency practices employed by quantitative political scientists, but must instead develop and follow their own.

We begin from the position that qualitative research is invaluable, generating knowledge that could not be produced through any other form of inquiry. Such research generally entails close engagement with one or more cases, producing thick, rich and open-ended data. These data are collected and used by scholars with a range of epistemological beliefs, producing a wide variety of interpretations and inferences.

For qualitative scholars who are comfortable with replication (i.e., the repetition of a research process or analysis in an attempt to reproduce its findings), the case for transparency makes itself. Without transparency there can be no replication. Yet even qualitative scholars who do not share a commitment to replication should value greater visibility of data and methods. For instance, those who believe that an important social scientific task is to encourage recognition of the extent and importance of cultural, historical and social diversity should acknowledge the value of transparency in permitting the record of actors speaking in their own voices to reach readers of social scientific texts. In short, the more sense scholars can make of authors' arguments and evidence, the better they can engage them, the more varied techniques they can use to evaluate and document their legitimacy, and the more scholars can enter the conversation.

Transparency in qualitative research needs to be achieved and evaluated in ways that are sensitive to the nature of qualitative data, how they are gathered, and how they are employed. As the list offered previously suggests (see footnote 3), qualitative data take on more varied forms than quantitative data, and are less-structured. In terms of data collection/generation, qualitative scholars very commonly gather their own data, rather than rely solely on a shared dataset. Evaluating the processes used to obtain data is a key element in assessing qualitative work – not least because those processes have a critical effect on the research product. With respect to employment, qualitative data are used in a range of research designs, including single case studies, small-n case studies, and various mixed-method designs. A variety of methods are used to analyze qualitative data (e.g., narratives, counterfactual analysis, process tracing, Qualitative Comparative Analysis, content analysis, ethnographic analysis), and different inferential structures underpin each method. These fundamental facets of qualitative research have implications for how transparency can and should be achieved.

These epistemological considerations are reinforced by the especially acute ethical and legal imperatives, and the sociological framing of transparency, in qualitative research. The two most important ethical and legal imperatives with which transparency can be in tension in qualitative research are human subject and copyright concerns. Sometimes data are collected in circumstances that require discretion to protect the rights and welfare of subjects. This will, quite properly, limit transparency. Moreover, many sources are not, in their entirety, in the public domain, and there are limitations on how they can be shared. As noted below, scholars should only make qualitative data (and information about the decisions and processes that produced them) available in ways which conform to these social and legal imperatives.

Sociologically, no amount of written guidance will result in changes in transparency practices unless scholars believe that methods and research goals about which they care are being preserved and improved. A separate set of guidelines for qualitative research helps to establish that the aim of transparency is to demonstrate the power of qualitative research designs, data-collection techniques, interpretative modes, and analytic methods. In other words, rather than tacitly encouraging changes to qualitative research practices, the goal of enhanced transparency in qualitative research is precisely to *preserve* and *deepen* existing qualitative research traditions, render current qualitative

research practices more accessible, and make clearer the tremendous value-added qualitative research already delivers.

In short, while transparency is a universal principle, for epistemological, ethical, and sociological reasons, its instantiation in qualitative research needs to conform to traditions specific to qualitative work.

Data Access

Clause 6.1 in the revised APSA *Ethics Guide* obliges a scholar who makes evidence-based claims in her published work to reference the data she used to make those claims. If the scholar generated or collected the data herself, then she should also make those data available or explain why she cannot.

What data should be referenced and/or made available, and how?

Researchers making evidence-based knowledge claims should clearly and completely reference the data on which they base their interpretations or their descriptive or causal inferences. Generally, these are the data the author explicitly cites to support those claims.

Referencing textual data requires a full and precise bibliographic citation including page numbers and any other information necessary for readers to locate the material cited and find within it the passage an author suggests is evidence for his claims. For primary archival sources, for instance, information about the archive and collection, and the number of the box in which the document was found should be included. For non-textual sources, information allowing an equivalent degree of precision should be included.[4] This information should be provided upon publication.

The new APSA standard entails a more stringent obligation for scholars who themselves generated or collected the data on which their evidence-based knowledge claims are based. Those scholars must, whenever possible, make those data available.[5] Later in this document, we discuss strategies for, and issues involved in, sharing qualitative data.

Sharing cited data is sufficient to meet the APSA standards. Nonetheless, for many qualitative researchers, cited data are often a small subset of the information collected and used in a research endeavor. As such, researchers are strongly encouraged to share data which are implicated in their research but not cited in their publication – for

---

[4] To give an example, when citing an audio tape, scholars might indicate the exact moment during the interview at which the cited material was mentioned (i.e., cite the time stamp), or might provide an extract of the recording and cite where it came from within the interview (e.g. this clip is six minutes in).
[5] As noted later, scholars using Active Citation to achieve transparency must provide substantial excerpts from the data sources underlying their claims (and ideally provide the actual data sources) no matter whether they generated or collected those data or other scholars did so. If the source or the relevant portion thereof cannot be provided for ethical or legal reasons, a summary or redaction must be offered.

instance, additional data used to generate the argument (rather than test it), or to infirm alternative interpretations and inferences.

What limitations might there be on making qualitative data available?

It is critically important that scholars sharing data comply with all legal and ethical obligations. As paragraph 6.4 of the *APSA Guide to Professional Ethics* notes, while it is incumbent upon researchers to accurately represent the research process and study participants' contributions, external constraints may require that they withhold data, for instance, in order to protect human subjects or to comply with legal restrictions.

Confidentiality and Human Subjects: If scholars have promised the individuals whom they involved in their research confidentiality, it is incumbent upon them not to reveal those subjects' identities. Personal identity can be disclosed both directly (for example, through divulging a participant's address, telephone number, age, sex, occupation, and/or geographic location) or indirectly (for example, by disclosing information about the person that, when linked with publicly available information, reveals his/her identity).

Data garnered from human subjects can often be shared legally and ethically if the appropriate informed consent is granted by project participants.[6] Where necessary, additional protective steps can be taken including guaranteeing confidentiality when soliciting informed consent;[7] employing anonymization strategies;[8] carefully controlling access to data; and/or requiring that special measures to protect confidential information be clearly specified in a data-use agreement signed by anyone who wishes to view or analyze the data.

Documentary Data: Sometimes the owners or licensors of data collected through non-interactive techniques—archives or non-governmental organizations, for instance—place limitations on their use or dispersion. Likewise, such materials sometimes have copyright restrictions. Scholars should make every attempt to explain the value of data-sharing to those from whom they acquire documentary data, and investigate to what degree, and which, copyright law applies.

---

[6] To be clear, there are instances in which a researcher who has obtained permission from a subject to share data should nonetheless not do so because, for example, the subject is not in a good position to evaluate the risk of information connected with him/her being made accessible, or the circumstances under which permission was granted have changed. Alternatively, the author may decide to impose additional safeguards not specified in the informed consent when sharing the data.

[7] When seeking informed consent researchers should secure permission for data sharing where possible, and should avoid including statements in consent forms that purposefully preclude data sharing beyond the researchers or team.

[8] For instance, on some occasions scholars will only be able to characterize their source ("a senior government official"), but will be able to attribute full quotations to him or her; on other occasions, they will be able to indicate that they consulted with particular people, but will not be able to attribute any specific information to them.

<u>Proprietary Data</u>: When research is based on proprietary data, authors should make available sufficient documentation so other scholars can evaluate their findings. Owners of proprietary data should be encouraged to provide access to bona fide researchers.

As the discussion of types of data 'friction' in this section makes clear, the exclusions and restrictions that can prevent authors from sharing the data that support their analytic claims are circumstantial, ethical and legal. Accordingly, where data cannot be shared, the author should clearly explain why not, and include as much information about those data as is ethically and legally possible, to help readers understand and evaluate the author's inferential and interpretive claims.

When should data be made available?

The APSA standards recognize that "Researchers who collect or generate data have the right to use those data first." A particular collection of data should be made available no more than one year after the earliest publication (either electronic or paper) of evidence-based statements made using that collection.

The APSA standards also recognize that journals and funding agencies may have different requirements (for instance, obliging researchers to make the data used in a book or article available prior to any publication). The one-year allowance specified by APSA does not alter any time limits established by journals and funding agencies.

Where and in what form should data be made available?

The best practice is for digital data (e.g. PDFs of documents, audio files, video files) to be made accessible online, at an established repository that can be discovered by standard Internet search engines. Standard and non-proprietary file formats are preferable, because they are more likely to remain accessible over time. For non-digital data, scholars should provide a metadata record identifying the source.

When deciding on a venue for making their data available, scholars should consider multiple desiderata. These include: the practices and rules of the publishing venue, the transaction cost for the reader of accessing the evidence in context, the potential storing venue's ability to make the data accessible to all interested persons, as well as to support annotation of citations (on which, more below), the likely durability of the venue (i.e., whether it has stable and long-term funding sources), the availability and quality of assistance with curation, and the cost to data users. [9]

Scholars who anticipate incurring incremental costs when preparing data for sharing (e.g., for anonymizing to protect confidential information) should consider building those costs into funding applications, and/or they may request reimbursement (perhaps

[9] Although university repositories will often meet these criteria, scholars are discouraged from hosting data themselves on local websites as such sites are notoriously unreliable. While doing so may be a good temporary measure, longer-term plans for storage in an established repository should be developed.

drawn from fees paid by researchers requesting to use shared data). Likewise, when distribution involves additional costs (e.g., for administration of special conditions of access to confidential information), data distributors may request reimbursement for the incremental costs of making data available (see Section 6.5 of the *Ethics Guide*).

What is a "persistent identifier"? Why should I get one? Where can I get one?

A persistent identifier is a permanent link to a publication, data collection, or unique metadata instance that points to (and records versioning of) a data collection on the Internet. The publisher of the resource agrees to maintain the link to keep it active. Over time the link behind the persistent identifier may be updated, but the identifier itself remains stable. There are several kinds of persistent identifiers (DOI, URN, Handle, etc.).

Persistent identifiers are "machine-actionable" and facilitate the harvesting of data references for online citation databases, like the Thomson-Reuters Data Citation Index. Scholars can easily track the impact of their data from citations in publications. An increasing number of journals are requiring persistent identifiers for data citations.

Persistent identifiers can be useful for keeping track of bodies of data. One way to obtain a persistent identifier for data is to deposit them in an established institutional or social science repository, for instance, members of Data-PASS (http://www.data-pass.org/).

Production Transparency

In order to achieve *production transparency*, researchers should provide comprehensive documentation and descriptive metadata detailing their project's empirical base, the context of data collection, and the procedures and protocols they used to access, select, collect, generate, and capture data. To offer three specific examples, authors should address basic issues of how documentary sources were selected or sampled, the terms under which interviews were granted, and how participant observation or ethnographic work was conducted.

Production transparency is a prerequisite for an author's data to be intelligible to other researchers. Providing information about decisions made and processes carried out in the course of collecting and generating data, selecting them for inclusion in published work, and presenting them makes it easier for other scholars to understand and interpret the data; allows them to assess whether those processes were carried out in an unbiased manner; and helps them to evaluate the validity of the claims made on the basis of the data.

The production transparency requirement is triggered when scholars themselves collected or generated the data that support their evidence-based claims. Accordingly, the same timetable and constraints that apply to making those data available apply to

production transparency in relation to those data. As noted previously, APSA allows scholars a one-year period for first use of data they collected and thus for describing the data-collection process.

If the data are subject to ethical or legal restrictions, it is likely that production transparency will be similarly constrained. Conforming production transparency to relevant limits helps to ensure that other scholars can evaluate or replicate authors' data-collection procedures legally and without threatening the privacy of human subjects.

Although documentation is often supplied in text files or spreadsheets, an advanced standard for documenting data (at the study level) in the social sciences is the Data Documentation Initiative (DDI). DDI is an XML markup standard designed for social science data. Since DDI is machine actionable, it can be used to create custom codebooks and to enable online search tools. A list of tools for creating DDI is available at the DDI Tools Registry (http://www.ddialliance.org/resources/tools). Original documents (e.g., technical reports, questionnaires, and showcards) can be submitted as text files or PDF/A.

Analytic Transparency

Achieving *analytic transparency* requires scholars to describe relevant aspects of the overall research process, detail the micro-connections between their data and claims (i.e., show how the specific evidence they cite supports those claims), and discuss how evidence was aggregated to support claims.

The APSA standard for analytic transparency prescribes no epistemology or methodology; it simply requires that authors be clear about the analytic processes they followed to derive claims from their data, and demonstrate how they followed the general rules that attend the interpretive or inferential approach they are using.

The Transparency Appendix and Active Citation

One way in which qualitative researchers can provide data access, achieve production transparency, and engage in analytic transparency, is by developing a transparency appendix to their published work. A transparency appendix typically consists of two elements: active citations and an overview section.

Active citations follow the format of traditional footnotes or endnotes, but are digitally augmented to include:
- a precise and complete reference and any additional information that scholars will need to locate the cited source and find the relevant information within it;
- excerpts from cited sources;
- the cited sources themselves if the author possesses them and is in a position to share them, and/or hyperlinks thereto;

- annotations that
    - explain how individual pieces of data, sources, citations, and facts were interpreted and why they were interpreted as they were;
    - illustrate precisely how those individual pieces support claims in the text;[10]
    - address any important interpretive ambiguities or counter-arguments;
    - explain how individual pieces of data aggregate to support broad interpretative and theoretical conclusions.

Because active citations follow the format of traditional footnotes or endnotes, they are ideally suited to elucidate particular inferences or interpretations in the author's text. Certain aspects of research that should be explained if transparency is to be achieved, however, do not comfortably attach themselves to a particular subsection of text or footnote. These matters are instead best dealt with holistically. When such overarching concerns cannot be addressed in the main text, authors should include a brief "overview" in the transparency appendix clarifying their overall research trajectory (e.g., how interpretations and hypotheses were generated and evaluated); outlining the data-generation process; and demonstrating how the analysis attends to the inferential/interpretive rules and structures that underlie the type of analysis the author is doing.

Information provided in a transparency appendix *supplements* rather than replaces or repeats information offered in the text and footnotes of a book or article: it supplies additional context and background to authors' research efforts, offering an opportunity for authors to describe the rigor and thoroughness of their research (and field research), and allowing other scholars to understand and evaluate the appropriateness of their use (and, where relevant, generation) of data. What is "appropriate" depends upon the interpretive or inferential structures implied by the author's underlying epistemology and employed in the type of qualitative research he or she is conducting.

With respect to data access, scholars using active citation provide excerpts from the data sources underlying their claims (and ideally provide the actual data sources). In terms of production transparency, authors who cannot provide basic information about data collection in the main text of their publications due to length-limitations can include additional information in an introductory overview.

As for analytic transparency, the traditional representation in qualitative research—elaboration of an argument in the text combined with a simple citation—is often inadequate to make the link between an argument and evidence apparent. The critical element in the evidence is often difficult to discern, and the evidence is often interpretable in multiple ways. Likewise, a passage in a source can often only be properly interpreted within a broader textual context. Moreover, abbreviated ("scientific" or endnote) footnote formats, shrinking word limits for published work, and

---

[10] The standards the author may seek to meet or rules he may follow when detailing these micro-connections can vary: they may include the scholar's interpretation of the relative persuasiveness and consistency of evidence, explicit process-tracing, discussing narrative plausibility, making arguments about the plausibility of counterfactuals, advancing a systematic scheme for weighting data, mixed method approaches, etc.. Analytic transparency requires only that scholars be consistent and transparent, so that the reader can follow how their overall conclusions follow from smaller-scale findings.

unfamiliarity with careful textual interpretation have rendered traditional journals (and even books) inhospitable forums for achieving rigorous analytic transparency.

In sum, the introductory overview component of a transparency appendix empowers authors to enhance readers' understanding of the context, design and conduct of research. Using active citation empowers authors to clarify the micro-connections between data, analysis, and conclusions. Both enhance the rigor and persuasiveness of qualitative research.

Publishers' Responsibilities

Journals, editors, and publishers should assist authors in complying with data access and research transparency guidelines.

Publishers should:
- inform authors of options for meeting data access and research transparency requirements;
- host scholars' cited sources and transparency appendices on line, or guide authors to online archives which will house these materials, and provide links from articles (at the level of the individual citation, if needed) to those materials;
- provide guidelines for bibliographic citation of data;
- include consistent and complete data citations in all publications.

Resources

- Corti, Louise. 2000. "Progress and Problems of Preserving and Providing Access to Qualitative Data for Social Research – The International Picture of an Emerging Culture." *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* 1 (December). http://www.qualitative-research.net/index.php/fqs/article/view/1019 (August 13, 2009).
- Corti, Louise. 2005. "Qualitative Archiving and Data Sharing: Extending the Reach and Impact of Qualitative Data." *IASSIST Quarterly* (Fall): 8-13.
- Swan, Alma and Sheridan Brown. 2008. "To Share or Not to Share: Publication and Quality Assurance of Research Data Outputs." A report commissioned by the Research Information Network. School of Electronics & Computer Science, University of Southampton. http://www.rin.ac.uk/files/Data%20publication%20report,%20main%20-%20final.pdf (August 13, 2009).
- UK Data Archive
  - Managing and Sharing Data: A Best Practice Guide for Researchers (http://data-archive.ac.uk/media/2894/managingsharing.pdf)
  - Create and Manage Data (http://data-archive.ac.uk/create-manage)
- UK Data Service
  - Advice and Training (http://ukdataservice.ac.uk/use-data/advice.aspx)
  - Prepare and Manage Data (http://ukdataservice.ac.uk/manage-data.aspx)

- Van den Eynden, Veerle and Louise Corti. 2009. "Tensions between data sharing and data protection in research with people." *SRA News* (May): 12-15.
- Wolf, Virginia A., Joan E. Sieber, Philip M. Steel, and Alvan O. Zarate. 2006. "Meeting the Challenge When Data Sharing is Required." *IRB: Ethics and Human Research* 28 (March-April): 10-15.

Appendix

Section 6 of the American Political Science Association's *Guide to Professional Ethics, Rights and Freedoms* as amended in October 2012:

"6. *Researchers have an ethical obligation to facilitate the evaluation of their evidence-based knowledge claims through data access, production transparency, and analytic transparency so that their work can be tested or replicated.*

6.1 Data access: Researchers making evidence-based knowledge claims should reference the data they used to make those claims. If these are data they themselves generated or collected, researchers should provide access to those data or explain why they cannot.

6.2 Production transparency: Researchers providing access to data they themselves generated or collected, should offer a full account of the procedures used to collect or generate the data.

6.3 Analytic transparency: Researchers making evidence-based knowledge claims should provide a full account of how they draw their analytic conclusions from the data, i.e., clearly explicate the links connecting data to conclusions.

6.4 Scholars may be exempted from Data Access and Production Transparency in order to (A) address well-founded privacy and confidentiality concerns, including abiding by relevant human subjects regulation; and/or (B) comply with relevant and applicable laws, including copyright. Decisions to withhold data and a full account of the procedures used to collect or generate them should be made in good faith and on reasonable grounds. Researchers must, however, exercise appropriate restraint in making claims as to the confidential nature of their sources, and resolve all reasonable doubts in favor of full disclosure.

6.5 Dependent upon how and where data are stored, access may involve additional costs to the requesting researcher.

6.6 Researchers who collect or generate data have the right to use those data first. Hence, scholars may postpone data access and production transparency for one year after publication of evidence-based knowledge claims relying on those data, or such period as may be specified by (1) the journal or press publishing the claims, or (2) the funding agency supporting the research through which the data were generated or collected.

6.7 Nothing in this section shall require researchers to transfer ownership or other proprietary rights they may have.

6.8 As citizens, researchers have an obligation to cooperate with grand juries, other law enforcement agencies, and institutional officials. Conversely, researchers also have a professional duty not to divulge the identity of confidential sources of information or data developed in the course of research, whether to governmental or non-governmental officials or bodies, even though in the present state of American law they run the risk of suffering an applicable penalty.

6.9 Where evidence-based knowledge claims are challenged, those challenges are to be specific rather than generalized or vague. Challengers are themselves in the status of authors in connection with the statements that they make, and therefore bear the same responsibilities regarding data access, production transparency, and analytic transparency as other authors."

APPENDIX B

DRAFT July 28, 2013

Guidelines for Data Access and Research Transparency for Quantitative Research in Political Science

The *APSA Guide to Professional Ethics, Rights and Freedoms* recognizes that

> 6. Researchers have an ethical obligation to facilitate the evaluation of their evidence-based knowledge claims through data access, production transparency, and analytic transparency so that their work can be tested or replicated.
>
>> 6.1 Data access: Researchers making evidence-based knowledge claims should reference the data they used to make those claims. If these are data they themselves generated or collected, researchers should provide access to those data or explain why they cannot.
>>
>> 6.2 Production transparency: Researchers providing access to data they themselves generated or collected, should offer a full account of the procedures used to collect or generate the data.
>>
>> 6.3 Analytic Transparency: Researchers making evidence-based knowledge claims should provide a full account of how they draw their analytic conclusions from the data, i.e., clearly explicate the links connecting data to conclusions.

Data Access, Production Transparency, and Analytic Transparency describe key stages of the research process. Data access is not sufficient without documentation of how data were prepared and how analysis was conducted. By meeting these requirements, researchers contribute to the credibility and legitimacy of Political Science.

While evidence comes in many forms, these guidelines refer primarily to numerical data that can be analyzed with quantitative and statistical methods.[1]

Data Access

What data should be accessible to other scholars?

> When an author makes evidence-based knowledge claims, all data required to replicate the results serving as evidence for statements and conclusions should be open to other scholars. Researchers who have generated or created their own data have an obligation to provide access to the data used in their analysis whenever possible. When the data were collected by others, an author is responsible for providing a clear path to the data through a full bibliographic

---

[1] A parallel set of guidelines is in preparation which are intended primarily for qualitative data.

citation. In both cases, the steps involved in deriving conclusions and inferences from data should be fully described.

Researchers are strongly encouraged to share data beyond those required for replication of published findings. It is particularly important for researchers to provide access to data used in the process of generating conclusions but not included in the final analysis. More generally, providing as much access as possible to existing data can increase its value and often attracts greater attention to the work of the people who produced it.

When should data access be provided?

The *APSA Guide to Professional Ethics* recognizes that "Researchers who collect or generate data have the right to use those data first." Data access should be provided no more than one year after public dissemination of evidence-based statements. Journals and funding agencies may have different requirements. Moreover, some funding agencies may require researchers to provide data access *prior to any publication*. Nothing in these guidelines should be read to contradict such requirements.

Where should data be made available?

Data should be made available online at an established repository or a website that can be discovered by standard Internet search engines. When deciding on a venue for making their data available, scholars should consider multiple desiderata, including the venue's ability to make the data available to all interested persons, the likely durability of the venue (does it have stable and long-term funding sources), the availability of assistance with curation, and the cost to data users.

How should data be made available?

All data should be accompanied by:
1. Documentation describing the data in full
2. A complete citation including a "persistent identifier," like "digital object identifiers" (DOIs).

Standard and non-proprietary formats are preferable, because they are more likely to remain accessible over time.

When distribution involves additional costs (e.g. for protection of confidential information), data distributors may request reimbursement for the incremental costs of making data available.

DRAFT   July 28, 2013

How do I share data that includes confidential information?

> As paragraph 6.4 of the *APSA Guide to Professional Ethics* notes, researchers may need to withhold access to data to protect subjects and comply with legal restrictions. However, secure methods of sharing confidential data are often available. When respondents might be re-identified by combining information in the data (e.g. age, sex, occupation, geographic location), a data use agreement specifying measures to protect confidential information can be required. Access may also be provided in a "data enclave," where information derived from the data can be reviewed before it is released.

What if I used proprietary data?

> When research is based on proprietary data, researchers should make available documentation that would allow other scholars to replicate their findings. Owners of proprietary data should be encouraged to provide access to all qualified researchers.

What is a "persistent identifier"? Why should I get one? Where can I get one?

> A persistent identifier is a permanent link to a publication or a dataset on the Internet. The publisher of the resource agrees to maintain the link to keep it active. Over time the link behind the persistent identifier may be updated, but the identifier itself remains stable and does not change. There are several kinds of persistent identifiers (DOI, URN, Handle, etc.)

> Persistent identifiers are "machine-actionable" and facilitate the harvesting of data references for online citations databases, like the Thomson-Reuters Data Citation Index. You will be able to easily track the impact of your data from citations in publications. An increasing number of journals are requiring persistent identifiers for data citations.

> The best way to obtain a persistent identifier is to deposit your data in an established repository. Social science repositories, like the members of Data-PASS (http://www.data-pass.org/), and institutional repositories assign persistent identifiers to their holdings. There are also agencies that will issue a persistent identifier to a website that you maintain yourself.

What are the obligations of scholars who use data collected by others?

> When the data were collected by others, an author is responsible for providing a full bibliographic citation in the same way that a publication or other scholarly product would be cited. Data citations should include author, title, date, and a persistent identifier (or other location information).

DRAFT   July 28, 2013

Production Transparency

Production transparency implies providing information about how original data were generated or collected, including a record of decisions the scholar made in the course of transforming their labor and capital into data points and similar recorded observations. In order for data to be understandable and effectively interpretable by other scholars, whether for replication or secondary analysis, they should be accompanied by comprehensive documentation and metadata detailing the context of data collection, and the processes employed to generate/collect the data. Production transparency should be thought of as a prerequisite for the content of one scholar's data to be truly accessible to other researchers.

What should documentation include about the overall research project?
- Principal Investigator
- Title
- Purpose of the study
- Scope of the study
- Study design
- Sample
- Mode of data collection
- Instruments used
- Weighting
- Response rates
- Funding source

What should the codebook provide about each variable?
- Variable description
- Instrument, question text, or computation formula
- Valid values and their meanings
- Cases to which this variable applies
- Methods for imputing missing values

How should I prepare documentation?

Although data producers often supply documentation in text files or spreadsheets, the standard for documentation in the social sciences is the Data Documentation Initiative (DDI). DDI is an XML markup standard designed for social science data. Since DDI is machine actionable, it can be used to create custom codebooks and to enable online search tools. A list of tools for creating DDI is available at the DDI Tools Registry.

DRAFT   July 28, 2013

Analytic Transparency

Scholars making evidence based knowledge claims should provide a full account of how they drew their conclusions, *clearly mapping the path from the data to the claims*. This path can be documented in many ways such as computer programs and scripts. Researchers should make available materials sufficient to allow others to reproduce their results. For example, when providing computer programs to satisfy an analytic transparency requirement, questions about sufficiency can be answered as follows:

Is the program that produced my tables enough?
> Transparency involves documenting all of the steps from the original data to the results supporting your conclusions.

I have lots of programs. Do I need to provide all of them?
> The best practice is to consolidate all data transformation and analysis steps in a single program. Program steps may be developed separately, but they should operate as an integrated workflow.

Publisher's Responsibilities

Journals, editors, and publishers should assist authors in complying with data access and research transparency guidelines.
Publishers should
- inform authors of options for meeting data access and research transparency requirements;
- verify that data and program code are accessible, when appropriate;
- provide guidelines for bibliographic citations of data;
- include consistent and complete data citations in all publications.

Resources
- Australia National University - ANU Data Management Manual, September 2010
- Columbia University, Center for International Earth Science Information Network (CIESIN) - Guide to Managing Geospatial Electronic Records, June 2005
- Council of European Social Science Data Archives (CESSDA) - Sharing Data Website
- Gary King - Data Sharing and Replication
- Inter-university Consortium for Political and Social Research (ICPSR) - Guide to Social Science Data Preparation and Archiving, 2009
- UK Data Archive - Managing and Sharing Data: A Best Practice Guide for Researchers, September 2009
- UK Data Archive - Create and Manage Data Website

Data Access and Research Transparency in the Qualitative Tradition

Colin Elman, *Syracuse University*

Diana Kapiszewski, *Georgetown University*

As an abstract idea, openness is difficult to oppose. Social scientists from every research tradition agree that scholars cannot just assert their conclusions, but must also share their evidentiary basis and explain how they were reached. Yet practice has not always followed this principle. Most forms of qualitative empirical inquiry have taken a minimalist approach to openness, providing only limited information about the research process, and little or no access to the data underpinning findings. What scholars do when conducting research, how they generate data, and how they make interpretations or draw inferences on the basis of those data, are rarely addressed at length in their published research. Even in book-length monographs which have an extended preface and footnotes, it can sometimes take considerable detective work to piece together a picture of how authors arrived at their conclusions.

There are multiple overlapping reasons why scholars might follow this minimalist approach. One root is a "craft" understanding of qualitative research. According to this view, only scholars who are as immersed in the background and detail of an author's cases, and as familiar with the bulk of the evidence as he or she is, are in a position to make an informed judgment about the conclusions. Given that those scholars are well-equipped to locate the relevant sources and understand their connection to claims, brief footnotes are deemed sufficient. In addition, because the discipline does not currently reward greater openness, scholars are reluctant to take on the logistical burden it seems to imply in the absence of a clear and substantial payoff. Other objections arise because qualitative data are often under constraint. For example, concerns about human subjects may limit sharing when it would impinge on the rights and welfare of study subjects. Copyright and other proprietary rights might also limit transparency as many sources may not be in the public domain.

We are not wholly convinced by these justifications for taking a minimalist approach to openness. We doubt that the level of detail most qualitative researchers currently provide about their data, and about the connections between their data and their conclusions, is sufficient for even well-informed interlocutors to fully appreciate their arguments. Given the transaction costs of "looking for themselves," even specialists end up relying on a trust norm and heuristic shortcuts.[1] Further, as the discipline becomes more aware of the benefits of openness, we anticipate that gatekeepers such as journals, funding agencies, and academic departments will increasingly implement openness as a merit criterion—meaning that for individual scholars, adopting transparency practices sooner rather than later is actually the rational route to take. Finally, while we recognize that human subjects and copyright concerns may place limitations on how much openness can occur, we believe that a great deal of information can be shared without crossing either ethical or legal boundaries.[2]

Starting from the position that qualitative research as it is currently practiced is a valuable enterprise, in this article, we investigate whether and how qualitative scholars can reveal more about the processes through which they generate, analyze, and deploy data. Our central message is that if qualitative scholars take a more self-conscious, deliberate, and expansive approach to data access and research transparency (DA-RT), they can demonstrate the power and rigor of their work more clearly and empower a much larger audience to understand and interpret their research on its own terms.

OPENNESS AS A "META STANDARD"

All rule-based social inquiry is based on three notions. First, scholarly communities hold shared and stable beliefs that research designed and conducted in particular ways—according to particular rules—is warranted to produce knowledge with certain characteristics. Second, both the conduct of social inquiry and the written products that represent its conclusions are designed to capture those characteristics. Finally, to possess those characteristics, research must be designed and conducted in accordance with those rules.

Thus when scholars claim to have explained, interpreted, predicted, or otherwise asserted knowledge about an aspect of the social world, the warrant for making that claim does not come solely from the data they collected and analyzed. It comes in part from theories of knowledge that argue that when data generation and analysis follow certain rules, such claims can be made. In turn, our evaluation of those claims is based on our assessment of whether scholars' research processes followed those rules.

Consider, for example, the connection between epistemology, method, and results in experimental research (Druckman, Green, Kuklinski, and Lupia 2011, 17). The departure point for such research is that randomized assignment addresses the fundamental problem of causal inference and that any difference in outcome between a control and treatment group will be, on

average, attributable to the treatment, that is, to the causal intervention. Hence research designs can be meaningfully divided into those that can establish causation and allow researchers to aver that they have done so (if the experiment is designed and conducted according to the relevant rules), and those that cannot.

Openness allows authors an opportunity to show they are conducting inquiry of a certain type, and thus can potentially make claims with particular characteristics. Derivatively, openness allows scholars to show that a particular piece of research was well designed and done correctly, and hence commands the virtues of those additional warrants. That is, our judgments as to the relative strength of research amount to assessments of whether it was produced in accordance with the relevant rules.

Foundational epistemologies, and the characteristics of their archetypal knowledge statements, differ across research communities. Nevertheless, as Lupia and Elman (2014, this symposium) note, at a general level, some common ground exists among them. The methodologies political scientists use to reach evidence-based conclusions *all* involve extracting

approach to qualitative inquiry, process tracing. We use this example to illustrate the underlying connections among epistemology, analytic method, data, and conclusions. Scholars who use process tracing engage with the social world, draw observations from it, generate data, conduct analyses, and deploy data to support their claims in particular ways. These contrast both with how scholars conducting large-N observational research carry out these tasks and how scholars operating in alternative sub-types of qualitative inquiry do so.[4] Given their varying foundations, these other approaches, by definition, take a different view of data and analysis. Correspondingly, their discussions of transparency would be different.

To be clear, our intention is not to provide a full account of process tracing itself. Rather, we hope to illustrate how greater transparency would make the technique's strengths more apparent, and focus critics on its real rather than its imaginary weaknesses.

To scholars unfamiliar with qualitative research, process tracing might seem like unsystematic soaking and poking

When scholars claim to have explained, interpreted, predicted, or otherwise asserted knowledge about an aspect of the social world, the warrant for making that claim does not come solely from the data they collected and analyzed. It comes in part from theories of knowledge that argue that when data generation and analysis follow certain rules, such claims can be made.

information from the social world, analyzing the resulting data, and reaching a conclusion based on that combination of the evidence and its analysis.

Hence, and as reflected in the revisions to APSA's ethics guide (Lupia and Elman, ibid.), openness requires all scholars to provide access to the data on which their conclusions were based and to clearly describe the analysis that produced those conclusions. Yet despite the universal applicability of openness as a meta standard, transparency is always instantiated locally.[3] That is, openness requires everyone to show their work, but what they show and how they show it varies. These differences are grounded in epistemic commitments and the rule-bound expectations of the tradition in which scholars operate.

RESEARCH TRANSPARENCY AND PROCESS TRACING: AN EXAMPLE

We have argued that how openness is achieved in a particular tradition depends on how its scholars generate and use data to gain inferential or interpretive leverage. Developing transparency practices for qualitative research thus requires understanding how (and why) the observations scholars draw from the social world are converted into data, how a subset of those data are used in support of their analytic claims, and how a subset of *those* data are selected for citation.

In this section, we discuss some issues that arise when pursuing transparency in the context of one within-case

for evidence that substantiates an empirical claim. In part, this misimpression is a function of process tracing being more sophisticated than its typical representation in published substantive applications suggests. As occurs with much qualitative work, the rigorous techniques that underlie process tracing often remain implicit in published scholarship (Mahoney 2012, 14, 21), rendering it vulnerable to critique. Transparency calls on qualitative scholars to make these techniques more explicit.

Among qualitative methodologists there is a widespread consensus that single pieces of data play a crucial role in process tracing.[5] Our key point is this: identifying those "diagnostic" data is an *analytic* procedure that scholars who use the technique should clearly describe in their work.

A qualitative datum, like a quantitative datum, is one among many: both are considered in the context of many other pieces of information. How the single piece of information relates to the group, however, is quite different in the two modes of analysis. A quantitative datum is one among many of the same thing, a comparable measure of the same characteristic in a sample. In process tracing, a qualitative datum is a single, unique piece of information that nevertheless gathers its meaning as part of the larger constellation of data in which it is embedded.

To be sure, as Andrew Moravcsik (2014, this symposium) persuasively argues, scholars should show the data they cite, and make clearer how those data support their claims.

Scholars should also, however, give some consideration to representing more of the body of material they consulted during research. Indeed, because qualitative scholars engage with the social world (rather than with a discrete data set), they inevitably encounter many more potentially relevant sources than they engage, draw many more observations from those sources than they convert into data, and generate more data than they cite in their research.

A cited datum has a complicated relationship with that larger uncited corpus. When scholars assert the importance of a qualitative datum drawn from a particular source by citing it, they are almost always implicitly making two statements about the *other* sources they consulted: observations drawn from those other sources contributed to giving the cited source its meaning, *and* none of these other sources contains more diagnostic information than the one cited. Thus what appears to be a wholly granular use of data in a process

To be clear, we are not suggesting that process tracing—or any other qualitative analytic technique—be done any differently than it is now. We simply encourage scholars to be clearer about what they did and to share more of the data that underlie their claims. In other words, we hope to expose the aspects of qualitative work that often remain invisible, thus helping qualitative scholars to convey the power of their research.

EVALUATING QUALITATIVE RESEARCH

Transparency is also a prerequisite for meaningful evaluation. As we noted earlier, claims and inferences produced by social inquiry are only valid to the extent that the work that produced them followed the rules of the relevant research tradition. Evaluating research thus entails assessing whether it was designed and conducted in ways that followed those rules.

In contrast to some other research traditions, qualitative scholars are only partially committed to replication as a mech-

Among qualitative methodologists there is a widespread consensus that single pieces of data play a crucial role in process tracing. Our key point is this: identifying those "diagnostic" data is an analytic *procedure that scholars who use the technique should clearly describe in their work.*

tracing study is actually somewhat holistic. Moreover, how a datum comes to be considered diagnostic is explicitly a product of research design. Hence, scholars writing process tracing accounts triage data to identify clearly delineated tests of necessity and sufficiency. It is because some observations are usable in hoop or smoking-gun tests that they are probative for the hypothesized explanations. That is, the logical requirements of the research design identify which data are diagnostic.[6]

Given the relevance of the large body of material consulted during research to a process tracing study, its partial representation in most such accounts runs counter to the nature and spirit of qualitative inquiry. It leads to an understatement of the contribution made by the broad set of data that qualitative scholars generate. It also compromises the persuasiveness of qualitative accounts. Traditionally cited data cannot fully represent the breadth of material used to draw inferences and arrive at conclusions in qualitative research. Of course it is impossible for scholars to show all of the data they used in an analysis; however, a project's evidentiary base—and the relationship between cited data and the broader set of data used in the analysis—can be more carefully described or at least better referenced.[7]

As APSA's ethics standards suggest, achieving transparency in a published process tracing account thus requires scholars do more on three fronts. They should more explicitly describe how they drew observations and generated data,[8] more precisely explain how they deployed those data and used process tracing to reach their conclusions, and share more of those data, than is currently the norm.

anism for establishing validity. This is especially true for evaluating data *generation*. Rarely does one qualitative scholar "redo" another's interviews, for instance—because sufficient information is not generally provided to do so and because the logistical burden is often prohibitively large. In addition, it would be unrealistic to expect the exact same information to be garnered given how contextual variables affect such interactions.[9] But, even without a commitment to replication, openness is necessary for readers to assess how scholars drew observations from sources, attached meaning to them, and identified them as analytically significant to their research.

For instance, readers can carefully assess whether the authors' data-generation techniques were aligned with the rules of inference and interpretation they were following. This alignment is a predicate requirement for the associated analytic methods to be successfully employed. Readers will also want to assess whether any given data-generation technique was used effectively. Hence, they might see what they can learn about how authors drew observations from sources, for instance, reading their interview transcripts to assess whether they asked leading questions or in some other way biased their interviews. Readers could also assess whether researchers engaged in triangulation, given the subjectivity inherent in many qualitative sources. At an extreme, readers could go to the research context in which authors worked and draw observations from different sources (e.g., interview a different set of relevant actors) and evaluate the consistency between those observations and those drawn from cited sources (i.e., engage in post hoc triangulation). As DA-RT practices develop, standards and norms for systematically generating data in qualitative work will become more explicit,

and scholars' data-generation techniques will be more easily evaluated against them.

With regard to evaluating the data *analysis* underlying the conclusions drawn in qualitative work, the most plausible standard is whether a reader could analyze the data cited as evidence and arrive at the same conclusions. Evaluating data analysis might entail assessing whether the operations authors performed on their data were appropriate (i.e., were dictated by the rules undergirding the form of analysis in which they were engaging). It might also involve assessing the micro-connections between individual pieces of (cited) data and descriptive/causal inferences or interpretation, again in view of the analytical methods the scholars were using, to determine whether the data support the inferences. Returning to the process tracing example, have the data cited been used in rigorous tests that confirm or disconfirm the competing hypotheses? As this discussion implies and as we suggested earlier, data sharing and analytic transparency are prerequisites for evaluating data analysis.

CONCLUSION: THE VALUE OF OPENNESS

Enhancing DA-RT has the potential to yield significant benefits for qualitative scholars and scholarship. DA-RT empowers researchers to provide a more complete description of the value-added their immersion in the social world provides. Although the details differ across research traditions, DA-RT

NOTES

1. For example, they might use extra-textual clues such as the number of interviews an author conducted or the time she spent in an archive to help assess whether she approached her task in a way that would have allowed her to generate useful data, effectively interpret them, and draw valid inferences.

2. For instance, soliciting appropriate informed consent (referencing data sharing) from project participants, effectively using anonymization strategies, controlling access to data, using precise data-use agreements, and sharing data in line with fair use standards can mitigate the concerns we highlighted previously. The bottom line, of course, is that scholars should only make qualitative data available in ways that conform to ethical and legal imperatives.

DA-RT allows qualitative scholars to demonstrate the power of their inquiry, offering an opportunity to address a central paradox: that scholars who value close engagement with the social world and generate rich, thick data rarely discuss the contours of that engagement, detail how they generated and deployed those data, or share the valuable fruits of their rigorous labor.

allows qualitative scholars to demonstrate the power of their inquiry, offering an opportunity to address a central paradox: that scholars who value close engagement with the social world and generate rich, thick data rarely discuss the contours of that engagement, detail how they generated and deployed those data, or share the valuable fruits of their rigorous labor.

This article focused on transparency techniques in the context of one within-case approach to qualitative research: process tracing. Other traditions have different understandings of how to be transparent about their research practices and grapple with their own difficult questions. Given these different concerns, the transparency conversation will be more productive if scholars from diverse research communities participate and begin to identify the levels and types of transparency with which they are comfortable and that are consistent with their modes of analysis. Thus the ideas we have proposed in this article represent one contribution to a broader ongoing debate that we hope will result in the gradual introduction and acceptance of transparency as a worthy goal, and useful standard, for social science research.

3. Sports and law provide analogues: all sports are governed by a meta standard of fair play, and all jurisdictions are governed by a meta standard of legality—yet what this means in practice differs from sport to sport, and jurisdiction to jurisdiction, because the rules differ.

4. Others include counterfactual analysis (e.g., Sekhon 2004), Qualitative Comparative Analysis and other forms of analysis with set-theoretic foundations (e.g., Goertz and Mahoney 2012; Grofman and Schneider 2009; Ragin 2000, 2008; Schneider and Wagemann 2010), and ethnographic analysis (e.g., Schatz 2009; Yanow and Schwartz-Shea 2006).

5. See Beach and Pedersen 2013; Bennett and Checkel forthcoming; Bennett 2010; Collier 2011; Mahoney 2012.

6. This is analogous to the increasing prevalence among quantitative scholars of the view that strong causal inference requires either data produced by an experiment or observational data from circumstances that mimic an experiment (i.e., a natural experiment).

7. Given the increasing use of web appendices and data archiving, the initial clause of this sentence may soon be inaccurate. Data and data collection practices can be described in the "overview" portion of a Transparency Appendix accompanying the published work of scholars engaging in active citation; see Moravcsik (2014, this symposium).

8. Indeed, qualitative scholars' penchant for generating their own data makes clearly describing how, when, where, and why data were generated—the "production transparency" aspect of DA-RT—particularly critical for research transparency in qualitative work.

9. Other forms of data collection may be more easily replicable. It is an open question whether archival research is more replicable than interviews. Assuming universal archival coverage is not possible, different researchers

are likely to take distinct trajectories through the materials and arrive at dissimilar interpretations of them.

REFERENCES

American Political Science Association. 2012. *A Guide to Professional Ethics in Political Science*, Second Edition. Washington, DC: American Political Science Association.

Beach, Derek, and Rasmus Brun Pedersen. 2013. *Process-Tracing Methods: Foundations and Guidelines*. Ann Arbor: University of Michigan Press.

Bennett, Andrew. 2010. "Process Tracing and Causal Inference." In *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, Second Edition, eds. Henry E. Brady and David Collier. Lanham, MD: Rowman and Littlefield.

Bennett, Andrew, and Jeffrey Checkel, eds. Forthcoming. *Process Tracing in the Social Sciences*. New York: Cambridge University Press.

Collier, David. 2011. "Understanding Process Tracing." *PS: Political Science and Politics* 44 (4): 823–30.

Druckman, James N., Donald P. Green, James H. Kuklinski, and Arthur Lupia. 2011. "Experiments: An Introduction to Core Concepts in Political Science." In *Cambridge Handbook of Experimental Political Science*, eds. James N. Druckman, Donald P. Green, James H. Kuklinski, and Arthur Lupia. New York: Cambridge University Press.

Goertz, Gary, and James Mahoney. 2012. *A Tale of Two Cultures: Qualitative and Quantitative Research in the Social Sciences*. Princeton, NJ: Princeton University Press.

Grofman, Bernie, and Carsten Schneider. 2009. "An Introduction to Crisp Set QCA, with a Comparison to Binary Logistic Regression." *Political Research Quarterly* 62 (4): 662–72.

Lupia, Arthur, and Colin Elman. 2014. "Openness in Political Science: Data Access and Research Transparency." *PS: Political Science and Politics* 47 (1): this issue.

Mahoney, James. 2012. "The Logic of Process Tracing Tests in the Social Sciences." *Sociological Methods & Research* 20 (10): 1–28.

Moravcsik, Andrew. 2014. "Transparency: The Revolution in Qualitative Research." *PS: Political Science and Politics* 47 (1): this issue.

Ragin, Charles. 2000. *Fuzzy Set Social Science*. Chicago: University of Chicago Press.

———. 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago: University of Chicago Press.

Schatz, Edward, ed. 2009. *Political Ethnography: What Immersion Contributes to the Study of Power*. Chicago: University of Chicago Press.

Schneider, Carsten, and Cladius Wagemann. 2010. "Standards of Good Practice in Qualitative Comparative Analysis (QCA) and Fuzzy-Sets." *Comparative Sociology* 9 (3): 397–418.

Sekhon, Jasjeet S. 2004. "Quality Meets Quantity: Case Studies, Conditional Probability, and Counterfactuals." *Perspectives on Politics* 2 (2): 281–93.

Yanow, Dvora, and Peregrine Schwartz-Shea. 2006. *Interpretation and Method: Empirical Research Methods and the Interpretive Turn*. Armonk, NY: ME Sharpe.

Transparency: The Revolution in Qualitative Research

Andrew Moravcsik, *Princeton University*

Qualitative political science, the use of textual evidence to reconstruct causal mechanisms across a limited number of cases, is currently undergoing a methodological revolution. Many qualitative scholars—whether they use traditional case-study analysis, analytic narrative, structured focused comparison, counterfactual analysis, process tracing, ethnographic and participant-observation, or other methods—now believe that the richness, rigor, and transparency of qualitative research ought to be fundamentally improved.[1]

The cornerstone of this methodological revolution is enhanced *research transparency:* the principle that every political scientist should make the essential components of his or her work visible to fellow scholars. Recognition of this principle recently led the American Political Science Association (APSA) to formally recommend higher transparency standards for qualitative (and quantitative) research (APSA 2012). The most broadly applicable tool for enhancing qualitative research transparency is *active citation*: a technologically enabled citation standard, according to which *any citation in a scholarly paper, article, or book chapter that supports a contestable empirical claim is hyperlinked to an excerpt from the original source and an annotation explaining how that excerpt supports the empirical claim, located in a "transparency appendix" attached to the document.* Active citation places the essential components of qualitative analysis—evidence, interpretation of evidence, and methodological selection criteria—just one click away from readers. This empowers them to engage more deeply with existing scholarship, not just as passive readers, but as active critics and authors of future research.

This article traces the changes, opportunities, and challenges posed for qualitative political science by the emerging disciplinary best practices of qualitative research transparency, particularly in the form of active citation. The first section defines research transparency in terms of three distinct dimensions and explains why it is a fundamental precondition for other advances in qualitative research. The second section explains precisely what the active citation is and why it is the most generally applicable and logistically convenient means to enhance qualitative research transparency.

WHY IS RESEARCH TRANSPARENCY ESSENTIAL?

Transparency is the cornerstone of social science. Academic discourse rests on the obligation of scholars to reveal to their colleagues the data, theory, and methodology on which their conclusions rest. Unless other scholars can examine evidence, parse the analysis, and understand the processes by which evidence and theories were chosen, why should they trust—and thus expend the time and effort to scrutinize, critique, debate, or extend—existing research?

Three Dimensions of Research Transparency

Research transparency has three dimensions: data, analytic, and production transparency. Recently APSA has officially recognized each of these three dimensions as creating professional obligations of ethical research practice (APSA 2012, 2013).[2] This section describes the three types of transparency and illustrates why each matters by pointing to weaknesses in current political science research.

Data transparency affords readers access to the evidence or data used to support empirical research claims. This permits readers to appreciate the richness and nuance of what sources actually say, assess precisely how they relate to broader claims, and evaluate whether they have been interpreted or analyzed correctly. Too often in qualitative political science today any effort to examine critical textual evidence ends in frustration. Authors rarely cite sources *verbatim* and almost never copiously enough to judge whether specific lines were cited in context. Those who would understand, critique, or extend existing work usually find it impractical to track down original sources. Incomplete or page-numberless citations are distressingly common: in a recent graduate seminar, my students found that even in the most highly praised mixed-method work, many sources (often 20% or more) could not be located by any means, including contacting the author. Even when sources can be identified, often the time, trouble, and translation difficulties required to get them impose prohibitive costs. Scouring university libraries, procuring books on inter-library loan, redoing field research to secure specialized publications or unpublished archival material, or reviewing an author's interviews or ethnographic observation notes is often impractical. Generally this means that only a few expert readers, and sometimes no one at all, has any inkling of what another scholar's qualitative data actually look like.

Analytic transparency assures readers access to information about data analysis: the precise interpretive process by which an author infers that evidence supports a specific descriptive, interpretive, or causal claim. Advancing a plausible argument about the precise meaning and reliability of a given piece of evidence requires a nuanced interpretation of it in a particular documentary, historical, strategic, cultural, and social context. Often this requires weighing alternative sources and interpretations and adjudicating ambiguities, tensions, contradictions, and synergies among them. Because this almost inevitably involves uncertain and potentially

contestable interpretations, analytical transparency requires that scholars provide an account of the basis on which they reached particular conclusions.

Whereas in the past analytic transparency remained at least theoretically feasible in published qualitative political science, which widely employed classic discursive footnotes, today it is largely precluded. Tighter word limits and the spread of so-called scientific citation forms designed for methodological approaches in which nonqualitative scholars only cite other secondary work, not actual evidence, make it nearly impossible for qualitative scholars to document claims properly.[3] Even when qualitative evidence is properly cited, it often remains on existing secondary sources, published primary material, or archival documents, interviews, ethnographic notes, and other primary evidence assembled by the author, it is almost invariably impossible for anyone except a few experts, often in different fields, to render even a *prima facie* assessment of how representative that data is. Very little scholarship explicitly mentions, let alone addresses in detail, the selection criteria for evidence. As regards method, qualitative analysts often discuss case selection, yet explicit discussions of specific methodological choices in how to design process tracing, counterfactual analysis, analytic narratives, ethnographic studies, or structured focused comparison are rare. Only with

Today qualitative political scientists seldom achieve a high degree of production transparency. Whether research rests on existing secondary sources, published primary material, or archival documents, interviews, ethnographic notes, and other primary evidence assembled by the author, it is almost invariably impossible for anyone except a few experts, often in different fields, to render even a prima facie *assessment of how representative that data is.*

obscure to readers precisely how descriptive, interpretive, and causal inferences were drawn, or what uncertainty attaches to each such analytical claims.[4] Only in exceptional cases are tensions among conflicting data sources addressed.

Production transparency grants readers access to information about the methods by which particular bodies of cited evidence, arguments, and methods were selected from among the full body of possible choices. Consider first evidence. Social scientific research results always face the concern that the particular observations—the measures, cases and sources—that an author has selected reveal only a subset of the data that could be relevant to the research question. This raises the danger of selection bias, which can occur due to conscious manipulation, unconscious "confirmation bias," or just plain sloppiness. This is a particular concern in data selection for qualitative case study work, in which scholars generally hand-pick sources, rather than using preassembled aggregate datasets. What, besides their conscience, prevents authors from cherry-picking evidence more likely to support a preferred description, interpretation, or causal theory? Similar concerns arise around the selection of specific theories, hypotheses, and methods: scholars must inevitably select certain frames, interpretations, theories, and methods for intensive attention, while setting others aside. (This conception of production transparency is broader than that in APSA standards, which only cover how data is selected (cf. APSA 2012, 2013).) Production transparency requires that scholars explain to the reader how such choices of evidence, theory, and method were made. At the very least, it gives readers a better awareness of the potential biases that a particular piece of research may contain. At most, the need for scholars to make this explicit will encourage and assist them to conduct less biased research.

Today qualitative political scientists seldom achieve a high degree of production transparency.[5] Whether research rests regard to the range of theories considered is there a common research practice (the "literature review") to provide some modest assurance that a proper range of explanations has been considered.

Data, analytic and process transparency concerns may seem picayune, yet they can be enormously important. Consider the example of Sebastian Rosato's recent book, *Europe United: Power Politics and the Making of the European Community* (Rosato 2011). Few scholarly works have received more scrutiny: it was published in a major book series, as an article and the subject of a symposium in *International Security*, and as a guide for current policy makers in *Foreign Policy*. Yet in all that time no one detected what two scholars who publish regularly in other disciplines quickly spotted: it establishes central theoretical claims by consistently cherry-picking sources and, more troublingly, by explicitly misreading or citing out of context many documents (often easily available secondary sources) to say precisely the opposite of what they unambiguously state. These biases drive the book's results. (Moravcsik 2013a; Lieshout 2012) Without greater transparency in political science, errors such as these are unlikely ever to be detected or debated. The truth is that we have no idea how reliable most existing qualitative political science really is.

Transparency: A Precondition for Improving Qualitative Research

Transparency is an essential foundation for rule-governed and intersubjectively valid social science research, in that it permits scholars to assess research and to speak to one another. It is also a precondition for any other advances in social science method, theory and data collection. As a recent draft APSA report on qualitative methods concludes (APSA 2013):

"Scholarly communities in the social sciences, natural sciences and evidence-based humanities can only exist if their members openly share evidence, results and arguments. Transparency allows those communities to recognize when research has been conducted rigorously, to distinguish between valid and invalid propositions, to better comprehend the subjective social understandings underlying different interpretations, to expand the number of participants in disciplinary conversations, and to achieve scientific progress."

In other words, any improvements in the quality of research presume transparency. Social scientists may assemble massive datasets and copious citations, deduce clever arguments from sophisticated theories, and use state-of-the-art methods, yet without transparent foundations, these serve no clear purpose, given that neither author nor reader can distinguish more and less compelling interpretations, accurate and inaccurate descriptions, or valid and invalid hypotheses.

Transparency is therefore not simply a precondition for assessing the quality of existing qualitative work, but also for encouraging and rewarding empirical, theoretical and methodological excellence in qualitative research. Without transparency, relatively little incentive exists to acquire new skills, collect better evidence, conduct superior data analysis, or render theory more accurate empirically. Aside from following methodological fashion, what external incentive does an individual scholar in a non-transparent setting have to improve research? Contemporary qualitative case study work in political science, which often lacks transparency, provides a useful illustration. Scholars today, especially graduate students and younger researchers, hesitate to invest in detailed linguistic training, deep area and functional expertise, intensive field work, and rigorous presentation and analysis of qualitative documentation. Scholarly debates and symposiums, journal reviews, and professional assessments of qualitative political science rarely assess its richness or rigor or question the empirical veracity of specific empirical claims (i.e., to what extent textual evidence actually supports theoretical claims). Instead, qualitative debates tend disproportionately to focus on abstract theoretical disagreements. So-called multi-method dissertations tend to invest years in careful and transparent formal models and statistical analysis, then in the final months quickly sketch in lower quality case studies.

For similar reasons, recent decades have seen a profusion of innovative but largely ignored methodological advice on how to better use sophisticated techniques of qualitative causal inference, including conventional narrative, counterfactual analysis, analytic narrative, structured focused comparison, process tracing, and ethnographic and participant observation.[6] Such techniques are surprisingly rarely used in empirical work, let alone in a sophisticated or innovative manner. One reason is that, without transparent evidence and data analysis, it is difficult to demonstrate to readers that empirical results are thereby more conclusive. The old adage that "one can prove anything with a case study" continues to be a self-fulfilling prophecy, even when methodologists have now conclusively shown it need not be so.

ACTIVE CITATION: THE CORE INSTRUMENT OF QUALITATIVE RESEARCH TRANSPARENCY

The revitalization of qualitative methods in recent years has focused on various tools for promoting research transparency. These include data archiving, qualitative data-basing, hyperlinks, traditional citation, and active citation. All are indispensable instruments for particular purposes. Data archiving is useful for preserving modest-sized collections of field data, such as interviews, ethnographic notes, and informal documents, although it suffers from major logistical, intellectual property, and human subject limitations.[7] Databases (using programs such as Access, Filemaker, or Atlas) can be extremely useful to manage qualitative data, particularly to support specific research designs where a moderate amount of evidence is analyzed to estimate and manipulate relatively few predefined variables across several cases, using intensive coding and mixed-method data analysis (Lieberman 2010). Yet they are relatively inflexible and have high up-front costs. Hyperlink citations to online sources, on the model of modern journalism and disciplines such as law and medicine, works for certain narrow applications of on-line research. Yet it fails to accommodate the majority of political science sources that are not found online or accessible under intellectual property law and human subject restrictions, as well as running up against the problem that links are surprisingly unstable over time. Currently, conventional footnotes remain the main instrument to assure research transparency, and they can work for simple cases. Yet as used in political science today, as we have seen, they often fail to provide a high degree of data, analytic or production transparency.
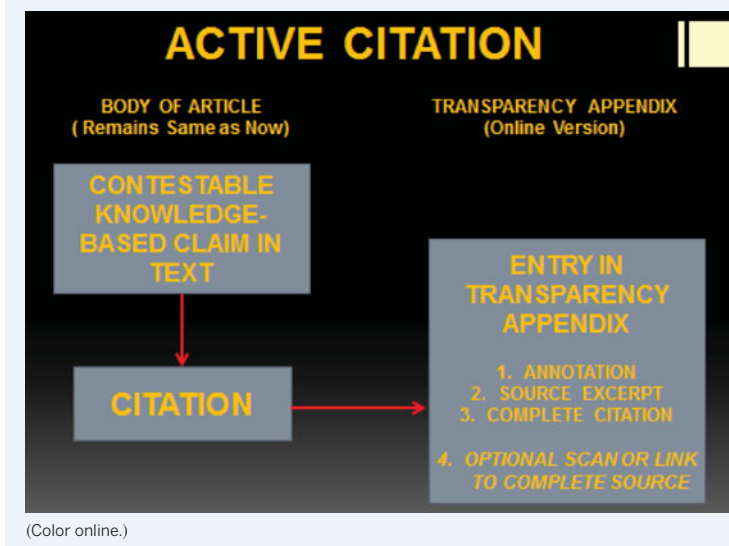
Active citation, by contrast, offers a general standard and format for presenting qualitative results that is far more general, flexible, logistically convenient, and epistemologically appropriate. Active citation envisages that *any empirical citation be hyperlinked to an annotated excerpt from the original source, which appears in a "transparency appendix" at the end of the paper, article, or book chapter*. The text of the article and the normal citation (footnote, endnote, or in-text citation) remain as they are now. Active citation requires only that the citation be complete and precise—a requirement already almost universally in place today, even if not always adhered to.

The distinctive quality of active citation involves the creation of a *transparency appendix* attached to the document. Each citation in the main text to a source that supports a contestable empirical claim is hyperlinked to a corresponding entry in the transparency appendix. Each entry contains four elements, three mandatory and one optional:

(1) a copy of the full citation
(2) an excerpt from the source, presumptively at least 50–100 words
(3) an annotation explaining how the source supports the claim being made
(4) optionally, an outside link to and/or a scan of the full source.

Figure 1 summarizes this format. While it is designed primarily with traditional textual sources in mind (e.g., primary

Figure 1.

(Color online.)

bolsters achieves a high degree of analytic transparency by including an annotation in the appendix entry, in which the author explains how and why the source supports the claim in the main text. This can also be used to elaborate any ambiguity, ambivalence, or uncertainty about that judgment, and to highlight the context (evidentiary, historical, cultural, social, or political) in which it was made. Active citation enhances procedural transparency by providing for entries in the transparency appendix—a specially dedicated one at the beginning and others as needed to support citations—to address issues of procedural transparency: how data, theories, and methods were chosen, and by what process the research was done. All of this is extremely convenient from the perspective of other scholars, because active citation connected all this information to the article, just one click away for the reader.

Active citation involves only a minor increase in workload for editors and authors. Only some citations need to be activated: not background references to literature reviews, theoretical debates or uncontested facts, but only those involving contestable empirical points. In a world in which scholars increasingly collect documents, conduct interviews, copy secondary sources, and keep records electronically, it is far less difficult than it once was to store, access, and input textual data—particularly if one anticipates active citation from the start. The standard remains deliberately flexible, so as not to create an undue burden on qualitative researchers who work under widely varying circumstances. The 50–100 word length of the excerpt, for example, is only a presumptive minimum. The actual length may be shorter or longer, and in extreme cases, with proper cause, may be replaced by a summary or omitted altogether. No one can be expected to cite verbatim text that cannot legally be excerpted, that is inconsistent with human subject or other institutional review board restrictions, or that imposes an undue logistical burden on the scholar. An article based on confidential interviews with Chinese military officers will probably not be as copiously sourced as one on nineteenth-century British documents. An interview that was not taped or transcribed cannot generally be cited *verbatim*. Such circumstances can be explained in the annotation. At the same time, more fortunate or ambitious scholars are able to reveal more detailed and extensive evidence, since the format retains the possibility for optionally inputting scanned documents or linking to online sources as a supplement to the transcribed excerpt. Over time, different research communities will likely develop distinctive practices and expectations concerning appropriate levels of documentation, reflecting their distinctive constraints.

The length of the annotation is subject to guidelines, but similarly remains ultimately at the discretion of the author. If the link between claim and evidence is obvious, straightforward, or trivial, one sentence should do. If the link is problematic and important, the author can and should explain it in detail. To minimize the logistical difficulties, free add-on

textual documents, published primary sources, interview or focus group data, oral histories, field notes, diaries and personal records, press clippings, pamphlets, and secondary sources), it is also compatible with photographs, maps, posters, art, audio clips, and other audiovisual material. The transparency index also contains special entries, one at the start and others if needed to support specific citations, which specifically address production transparency: how data, theories, and methods were chosen and by what process the research was done. For journals, the appendix would lie outside of conventional word limits and in most cases, one suspects, would only appear in the online version of the journal. To see an example of an active citation, pertaining to remarks of Thaddeus Stevens in Steven Spielberg's recent film, *Lincoln*, click on the "activated" citation to this sentence or, if you are reading this in hard copy, go to the link listed in the corresponding reference at the back of the article (Moravcsik 2013b).

Active citation enhances data, analytic, and production transparency with relatively modest changes and logistical demands, as compared to current scholarly practices. In most respects it is a conservative reform, involving only proper application of current standards, a modest extension of traditional practices, or the adoption of best practices from other disciplines. The core notion of active citation, namely that qualitative scholars must provide greater access to data and analysis, is already commonplace in fields such as law, public policy, journalism, classical philology, education and history. The use of electronic resources and appendices to achieve transparency is a staple of natural sciences, medicine, and law. The use of these techniques as "best practices" in other fields suggests that its demands are not logistically onerous or unreasonable.

Active citation straightforwardly bolsters data transparency by providing a brief excerpt of the source material, presumptively 50–100 words. If intellectual property, human subject, and logistical considerations permit, a scan or outside link can also take the reader to the full source. Active citation

software to MSWord (and eventually to LaTeX) is in development, which will automate the creation of the transparency appendix and its entries. Demonstration protocols already exist in current software (Moravcsik 2012b).

From the perspective of editors and publishers, active citation can be implemented with only minimal changes to existing paper journal and book publication formats. The main body of a scholarly work, including citation forms, remains unchanged, with the small exceptions of hyperlinks (in online versions) and full citations (already required but not always provided). Even article word limits—which have shrunk over

notion not just if one believes political science ought to be "replicable" in the strict sense that any given body of evidence can only be properly interpreted in a single way. No matter what their epistemology, anyone who seeks to generalize about politics should embrace efforts to multiply the variety and subtlety of case study evidence, and to increase the observations from which social scientists can draw descriptive and causal inferences. Those sympathetic to traditional history, interpretivist analysis, constructivist theory or critical social science may have even more reason to welcome enhanced transparency. Such scholars are keenly aware that

Active citation involves only a minor increase in workload for editors and authors. Only some citations need to be activated: not background references to literature reviews, theoretical debates or uncontested facts, but only those involving contestable empirical points.

time in a way biased against qualitative scholarship—remain unchanged, with excess content appearing in appendices that resemble existing formats for quantitative or formal appendices, or "supplementary materials" appendices in natural science journals. Active citation can be used in parallel ways with unpublished papers, working papers, online publications, published articles, book chapters, or any other any scholarly form. This means that the transparency appendix could be part of the journal article submission, and would be subject to review—thus eliminating the problem (which often arises with respect to quantitative work) of imperfect *ex post* enforcement of transparency rules. Journals may decide whether to publish the transparency appendix in the hard-copy version or only in online versions.

Active citation is currently being realized. The National Science Foundation (NSF) is funding a demonstration project, which has commissioned several dozen of the leading scholars in international relations and comparative politics to retrofit classic articles and forthcoming work to the active citation format. They will appear for public viewing on the Qualitative Data Repository at Syracuse University. The team running this program has developed a detailed set of guidelines covering the details of how to construct active citations. NSF funding is also being used to fund computer scientists, who are currently completing software add-ons to popular word-processing programs to automate the creation of a transparency appendix and the entry of individual entries. Conferences and workshops have been held on specific intellectual property, logistics, and human subject concerns, and discussions are being held with major journals and publishers, some of whom are working toward adopting the standard. Several journals are considering adoption.

CONCLUSION

Regardless of their approaches to studying politics, all scholars should embrace the obligation research transparency creates to share with their colleagues critical evidence, interpretive judgments, and procedural decisions. This is an attractive

research conclusions often rest on subtle interpretive judgments drawn from ambiguous evidence about political choices made in specific social, cultural, gendered or institutional circumstances (Geertz 1973). Yet most readers of qualitative scholarship today find it difficult to "get inside the heads" of the individuals and groups that other political scientists study. The perceptions, beliefs, interests, cultural frames, identities, deliberative processes, and (often non-rational) strategic choices of those individuals and groups are more often assumed, asserted or implied than actually portrayed empirically. Active citation offers immediate access to the textual record, thereby permitting those real-world individuals and groups to speak directly to readers in their own voices. This can convey a more vivid and immediate sense of politics as it is actually lived, as well as a better understanding of why they act as they do. No one has an interest in anonymous and context-less political science.

Active citation also vindicates a deeper insight of traditional historical and non-positivist epistemologies, namely that comprehending political life is in many ways an essentially interpretive enterprise, one that requires that readers recognize and engage not just the world-views of the human subjects who are being analyzed, but also those of the scholars who conduct the analysis. By revealing these worldviews through enhanced data, analytical and production transparency, active citation can help convey not just a richer and more nuanced impression of political life, but a more accurate understanding of what real actors perceive as being at stake in it, why they make the decisions they do, why scholars who analyze those decisions disagree, and how their colleagues could generate new research insights in the future. For these reasons, qualitative research transparency is a standard that should bring together political scientists of all epistemological and theoretical persuasions.

Goertz, Alexander Lanowska, Jack Levy, Arthur "Skip" Lupia, James Mahoney, John Owen, and Elizabeth Saunders for comments. ∎

NOTES

1. Recent works in the qualitative revival include King, Keohane, and Verba 1994, Van Evera 1997, Elman and Elman 2001, Gaddis 2002, Brady and Collier 2004, Bennett and George 2005, Bennett and Elman 2006, Bennett and Elman 2007, Trachtenberg 2009, Klotz and Prakash 2009, Moravcsik 2010, 2012a, 2013a, forthcoming, Goertz and Mahoney 2012, Beach and Pedersen 2013, Mosley 2013.

2. APSA documents refer to the first goal as "data access," which I have changed to "data transparency."

3. Exceptions include Christensen 2011 and Snyder 1989.

4. For admirably transparent analysis, see Khong 1992.

5. For an exception, see Saunders 2011.

6. See citations in footnote 2.

7. For further discussion of these options, see Moravcsik, forthcoming.

REFERENCES

American Political Science Association (APSA). 2012. Guide to Professional Ethics in Political Science. (Changes adopted by the Committee on Professional Ethics, Rights, and Freedoms, July.) Available at: https://www.apsanet.org/content_2483.cfm.

———. 2013. "Guidelines for Data Access and Research Transparency for Qualitative Research in Political Science." August 7. Available at: http://www.apsanet.org/media/PDFs/QUAL%20DART%20Guidelines%20for%20Data%20Access%20and%20Research%20Transparency%20August%207%202013.pdf.

Beach, Derek, and R. B. Pedersen. 2013. Process-Tracing Methods: Foundations and Guidelines. Ann Arbor: University of Michigan Press.

Bennett, Andrew, and Colin Elman. 2006. "Qualitative Research: Recent Developments in Case Study Methods." Annual Review of Political Science 9: 455–76.

———. 2007. "Case Study Methods in the International Relations Subfield." Comparative Political Studies 40 (2): 170–95.

Bennett, Andrew, and Alexander George. 2005. Case Studies and Theory Development in the Social Sciences. Cambridge, MA: MIT Press.

Brady, Henry, and David Collier, eds. 2004. Rethinking Social Inquiry: Diverse Tools, Shared Standards. Lanham, MD: Rowman and Littlefield.

Christensen, T. J. 2011. Worse Than a Monolith: Alliance Politics and Problems of Coercive Diplomacy in Asia. Princeton, NJ: Princeton University Press.

Elman, Colin, and M. F. Elman. 2001. Bridges and Boundaries: Historians, Political Scientists, and the Study of International Relations. Cambridge, MA: The MIT Press.

Gaddis, John. L. 2002. The Landscape of History: How Historians Map the Past. Oxford: Oxford University Press

Geertz, Clifford. 1973. "Thick Description: Toward an Interpretive Theory of Culture." In The Interpretation of Cultures, ed. C. Geertz, 3–30. New York: Basic Books.

Goertz, Gary, and James Mahoney. 2012. A Tale of Two Cultures: Qualitative and Quantitative Research in the Social Sciences Princeton, NJ: Princeton University Press.

Khong, Yuen Foong. 1992. Analogies at War: Korea, Munich, Dien Bien Phu, and the Vietnam Decisions of 1965. Princeton, NJ: Princeton University Press.

King, Gary, Robert O. Keohane, and Sidney Verba. 1994. Designing Social Inquiry: Scientific Inference in Qualitative Research. Princeton, NJ: Princeton University Press.

Klotz, Audie, and Deepa Prakash. 2009. Qualitative Methods in International Relations: A Pluralist Guide. New York: Palgrave Macmillan.

Lieberman, Evan S. 2010. "Bridging the Qualitative-Quantitative Divide: Best Practices in the Development of Historically Oriented Replication Databases." Annual Review of Political Science 13: 37–59.

Lieshout, Robert. 2012. "Europe United: Power Politics and the Making of the European Community by Sebastian Rosato (Review)." Journal of Cold War Studies 14: 234–37.

Moravcsik, Andrew. 2010. "Active Citation: A Precondition for Replicable Qualitative Research." PS: Political Science and Politics 43: 29–35.

———. 2012a. "Active Citation and Qualitative Political Science." Qualitative & Multi-Method Research 10: 33–37.

———. 2012b. "Methodological Memo: How to Create Active Citations Using MSWord" (Unpublished Memorandum, Princeton University, March 2012). Available at: http://www.princeton.edu/~amoravcs/library/active_citations.docx.

———. 2013a. "Did Power Politics Cause European Integration? Realist Theory Meets Qualitative Methods." Security Studies 22: 1–18.

———. 2013b. "An Example of an Active Citation: James Scovel on Thaddeus Stevens' Famous Quotation" (Methodological Memo, Princeton 2013). Available at: http://www.princeton.edu/~amoravcs/library/example.docx.

———. forthcoming. "One Click Away: The Transparency Revolution in Qualitative International Relations." Security Studies.

Mosley, Layna. 2013. Interview Research in Political Science. Ithaca, NY: Cornell University Press.

Rosato, Sebastian. 2011. Europe United: Power Politics and the Making of the European Community. Ithaca, NY: Cornell University Press.

Saunders, Elizabeth. 2011. Leaders at War: How Presidents Shape Military Interventions. Ithaca, NY: Cornell University Press.

Snyder, Jack. 1989. The Ideology of the Offensive: Military Decision Making and the Disasters of 1914. Ithaca, NY: Cornell University Press.

Trachtenberg, Marc. 2009. The Craft of International History: A Guide to Method. Princeton, NJ: Princeton University Press.

Van Evera, Steven. 1997. Guide to Methods for Students of Political Sciences. Ithaca, NY: Cornell University Press.

Data Access and Research Transparency in the Quantitative Tradition

Arthur Lupia, *University of Michigan*
George Alter, *University of Michigan*

The number of people conducting scientific analyses and the number of topics being studied are higher than ever.[1] At the same time, there are questions about the public value of social scientific endeavors, particularly of federally funded quantitative research (Prewitt 2013). In this article, we contend that data access and research transparency are essential to the public value of the enterprise as a whole and to the credibility of the growing number of individuals who conduct such research (also see Esterling 2013).

By quantitative research, we mean work that includes survey research, experiments, and mathematical and computerized models of dynamic processes. In this work, scholars convert attributes of observations and events into symbols. These symbols are joined with a grammar—typically a set of logical rules from mathematics or statistics—to form an inferential language. The resulting language of quantitative social science can produce a more precise description of concepts and relationships than ordinary language. Quantitative conclusions about causal relationships and existential propositions are often offered as direct numerical expressions or as exact functional forms.

With this precision, however, comes a potentially important limitation. Information can be lost when converting observations into symbols and interpreting these symbols via logical rules. When scholars fail to document—and make available to others—information about how they selected cases to study, particular attributes of the cases on which to focus, specific ways of converting these attributes into numbers, and the choice of certain types of mathematics or statistics to convert these numbers into knowledge claims, the meaning and value of quantitative social science knowledge claims becomes increasingly uncertain.

To more effectively and rigorously answer questions about the value of quantitative social science, it is imperative that those of us who conduct such research take actions that reinforce its credibility and make it easier for others to interpret our findings accurately. This means sharing our data whenever possible. It also means making available a complete description of the steps that we used to convert data about the social world into quantitative claims about how it does and does not work. Such commitments will not only help others more accurately assess our claims about individual events but also increase the extent to which others will view as credible our attempts to draw generalizations about people, policies, and institutions from a series of numerical simplifications and logical transformations.

With such imperatives in mind, political science, like other disciplines, is seeking to increase its credibility by improving procedural transparency and data sharing. Supporting efforts to increase data sharing and research transparency is the fact that technical barriers to such openness are falling quickly. Data archives, for example, are becoming more numerous and archivists have found multiple ways to make them easier to use. Viewed from a technical perspective, depositing one's own data and documents and accessing others data and documentation has never been easier. Old explanations for why scholars need not share data or procedural information are becoming more difficult to support. Indeed, younger generations, who do not remember life before the Internet, expect greater transparency of all kinds (Pew Research Center 2013). At the same time, we recognize that there have often been few incentives for taking the time to document one's procedures or to share one's data. Unless greater incentives for sharing data and publicizing analytic procedures emerge, it is difficult to expect old patterns to change.

In what follows, we describe current and future activities that support greater data sharing and research transparency. We focus in particular on several efforts to make data sharing and research transparency more rewarding for individual investigators and larger research collaboratives. We contend that these and other efforts can help individual political scientists, and the discipline, more effectively demonstrate the evidentiary and intellectual foundations of their insights. In so doing, this new emphasis on clarifying the evidentiary and logical foundations of one's knowledge claims can increase the credibility of individual research projects, reduce uncertainty about the meaning of social scientific findings, and increase the value of quantitative social science to multiple constituencies.

DATA SHARING

Many scientific disciplines are having broad conversations about data sharing. While there is much interest in the topic among political scientists today, leading figures in our discipline have long been engaged this topic. Warren Miller, a principal architect of the American National Election Studies (ANES), also founded the Inter-university Consortium for Political and Social Research (ICPSR), the world's largest social

science data archive. When seeking funding for data sharing, he (1962) argued that: "The sporadic work of individuals with diverse interests and little or no association with the work of predecessors or contemporary colleagues often produces worthwhile results but through the Consortium the power of extended and cumulative programs of research can also be realized." Miller also provided an example for others to follow by sharing ANES data from early in the project's inception. Today, thanks to the work of Miller and others, many data archives exist and the emergence of the Internet makes accessing and sharing data more feasible to more people.

Despite these changes, data sharing remains a contentious issue. Despite widespread acceptance of replication as a nominal virtue in quantitative social science, as Hamlet might say, "it is a custom more honour'd in the breach than the observance." Many scholars do not share their data (Acord and Harley 2012). Nonsharing scholars sometimes describe the work that they have put into collecting such data and opine that sharing data limits their ability to claim credit for their labor. Other scholars face confidentiality issues and believe that sharing data would open them to legal risk or minimize their oppor-

the time of publication. Other journals and professional associations allow an additional "embargo period" after publication. Because the time from submission to publication is usually more than a year and often much longer and since the time from collection to submission also entails periods of multiple months or years, either rule provides several years for authors to exploit their data.

Moreover, journals typically require only the data needed to support empirical claims in a specific publication, so-called "replication data." At present, authors are not expected to release aspects of a data set not supporting a publication.[2] Such practices represent a compromise between the opportunities for data collectors to reap the fruits of their labor and the broader benefits that can come from making that data available to others.

Another means of providing credit to scholars who create and share valuable data is to develop better data citation practices (Altman and King 2007). For centuries, scholars have been expected to cite the evidence, theories, and conclusions on which their own research builds. Notwithstanding this long tradition, social science editors and publishers have been slow

For centuries, scholars have been expected to cite the evidence, theories, and conclusions on which their own research builds. Notwithstanding this long tradition, social science editors and publishers have been slow to recognize that data are intellectual products for which citation should be required (Mooney 2011).

tunities for future data gathering. Yet other scholars want to share their data but do not know how to do so, or begrudge the time required to prepare data for sharing. In the remainder of this section, we address each of the barriers to data sharing in turn.

The first barrier pertains to claiming credit. Scholars who have invested a great deal of time and capital into collecting data want to reap benefits from their research. Some researchers resent the possibility that they may be "scooped" or even contradicted with their own data.

Scientific organizations are developing and implementing a number of creative ways to manage such difficulties. Underlying these endeavors are two premises. The first premise is that data that are shared widely can produce a larger number of meaningful inferences. Hence, sharing data is a public good that should be rewarded. The second premise is that data access facilitates replication that can be used to evaluate the truth-value and meaning of empirical claims. To this end, a growing number of journals are requiring authors who make empirical claims to make available the data from which such claims were derived.

At the same time, many professional associations, funding agencies, and journals acknowledge the need to balance the incentives of data collectors with the benefits to science and society from data sharing. One way that such a balance is sought is to offer a reasonable time to analyze data before making it available to others. Among social science journals, the most common requirement is for data to be available at

to recognize that data are intellectual products for which citation should be required (Mooney 2011). This pattern is changing. Numerous organizations actively promote data citation (see DataCite 2013; Data Preservation Alliance for the Social Sciences 2013; International Association for Social Science Information Services and Technology 2013; International Council for Science: Committee on Data for Science and Technology 2013; Inter-university Consortium for Political and Social Research 2013).

Funding agencies are also using data citation as a way to demonstrate the impact of their data collection investments. For example, in 2012, the ANES began to require people who download data from its website to sign an agreement to formally cite the ANES in any and all intellectual products they derive from its data. ICPSR has also expanded its ability to facilitate citation (http://www.icpsr.umich.edu/icpsrweb/ICPSR/citations/index.jsp). In 2012, Thomson Reuters unveiled a citation index for data, as a new feature in its Web of Knowledge platform (http://wokinfo.com/products_tools/multidisciplinary/dci/). With improvements in data citation, scholars who produce data that is used by others can expect to be cited for that work. Given the value that citation counts have for tenure and promotion reviews, better data citation norms should increase individual incentives to share data.

To this end, it should also be noted that a growing number of inquiries into citation counts find that "papers with publicly available datasets receive a higher number of citations than

similar studies without available data" (Piwowar and Vision 2013, 1). While we are not aware of such comprehensive studies being conducted in the social sciences, a study of more than 10,000 published studies in the gene expression literature funds that "studies that made data available in a public repository received 9% (95% confidence interval: 5% to 13%) more citations than similar studies for which data was not made available" (Piwowar and Vision 2013, 7). Dorch (2012) finds larger effects, albeit on a smaller sample of astrophysics articles.

The second barrier points to human subject harms and legal risks associated with some forms of data sharing. Protecting the privacy of research subjects is of paramount impor-

detailed plans for data security and other measures to prevent disclosure of confidential information. Highly sensitive data can be shared in controlled environments, like the Research Data Centers operated by the Census Bureau, where outputs can be screened for disclosure risks. Indeed, many organizations now operate remote execution systems or "virtual data enclaves," which allow researchers to conduct such analyses without having direct data access.

Although a wide range of measures have been developed to protect confidential information from research subjects, restrictions imposed by private organizations are a growing concern. Many corporations are amassing vast quantities of data. Because these data are considered commercial assets,

> *With improvements in data citation, scholars who produce data that is used by others can expect to be cited for that work. Given the value that citation counts have for tenure and promotion reviews, better data citation norms should increase individual incentives to share data.*

tance to the scientific community. Violations of confidentiality can damage public confidence in research and undermine the cooperation of subjects on which researchers depend. So the challenge becomes how to protect those rights while accruing, as much as possible, the individual and social benefits of data sharing.

The process of reconciling subject protection with data sharing begins with informed consent. Consent forms should promise to protect confidential information to the maximum extent allowed by law, but they need not exclude sharing data with other researchers.[3] After data are collected, a variety of techniques can minimize disclosure risks. Data "masking," for example, refers to techniques that modify data to prevent subject identification (see, e.g., Rubin 1993). Masking procedures include anatomizing (Xiao and Tao 2006), permuting (Zhang et al. 2007) or perturbing (Adam and Worthmann 1989) cells in a data matrix in ways that preserve the aggregate properties in which analysts are interested while decoupling identifying information from the data. When implemented successfully, these techniques allow analyses to be identical to what they would have been had individual cells not been altered. The ability to achieve such outcomes depends on relationships among properties of the data, how the cells are perturbed, and the kinds of analyses that individuals want to run. In a dataset with hundreds of variables, for example, it is typically impossible to implement a masking algorithm in ways that preserve all possible statistical relationships among all variables. If, however, a relatively limited set of relationships are of interest, successful masking possibilities emerge (Rubin 1993). Such techniques are expanding circumstances in which data can be shared while simultaneously protecting the privacy of individual respondents (Fung et al. 2010).

Other subject protection measures are also available (National Research Council 2003, 2005). Researchers can be required to sign data use agreements in which they provide

researchers are often required to sign nondisclosure agreements that prevent them from sharing the data with others. The *American Economic Review* exempts authors from sharing proprietary data but asks them to inform others how the data may be obtained (American Economic Association 2013). Political science journals should consider analogous policies.

Regarding the third concern, planning and good data management practices can reduce burdens often associated with data sharing. Many universities employ "data librarians" who offer assistance with data management planning. Data archives, including the six partners in the Data Preservation Alliance for the Social Sciences (http://www.data-pass.org/), are also available to offer advice and assistance to a wide range of scholars (see also ICPSR 2012). Professional archives can also help scholars document their work so that it remains accessible and functional for scholars who seek accurate interpretations of shared data. Such practices can also benefit data producers because well-designed documentation of data and research "workflow" can reduce the time needed to respond when a journal issues a "revise and resubmit" (Long 2009).

Changes in data citation practices, the possibility that articles associated with data sharing are more often cited, statistical masking, and professional archiving services are all factors that make data sharing more rewarding and feasible. To the extent that scholars come to formally cite all data that they use, quantitative social scientists will not only find data sharing more rewarding, but they will also be able to benefit from the data that others are now sharing. If scholars further commit to making data accessible and following the best practices of professional archivists, they and others can benefit for years to come from data that has already been created.

RESEARCH TRANSPARENCY

Sharing data does not provide all of the information about a quantitative analysis that can advance science and benefit soci-

ety. When assessing the meaning of quantitative claims, audiences often want to understand the decisions and actions that produced the claim. This follows because the meaning of a conclusion depends on the premises and practices from which it was derived.

In 2012, the APSA responded to calls for greater transparency by revising its "Guide to Professional Ethics, Rights, and Freedoms." The guide now states that "Researchers have an ethical obligation to facilitate the evaluation of their evidence-based knowledge claims through data access, production transparency, and analytic transparency so that their work can be tested or replicated." The distinction of production and ana-

nonresponse. Each of these decisions can affect the meaning of specific data points as well as the aggregate conclusions drawn from survey data. For survey analysts, transparency includes descriptions of how variables were coded and how analysts chose among different inferential methods and model specifications.

Today, information on the data production and analytic decisions that underlie many published works in political science is unavailable. This is one reason that many graduate courses in political science are unsuccessful in their attempts to replicate published empirical claims. Even when students have access to the same data as the original researchers (e.g.,

Despite the disappointing record for data sharing in some quantitative communities, promising signs indicate that research transparency is being taken more seriously in important areas of political science.

lytic transparency in the revision follows from definitions of Lupia and Elman (2010).

> Production transparency implies providing information about how the data were generated or collected, including a record of decisions the scholar made in the course of transforming their labor and capital into data points and similar recorded observations. In order for data to be understandable and effectively interpretable by other scholars, whether for replication or secondary analysis, they should be accompanied by comprehensive documentation and metadata detailing the context of data collection, and the processes employed to generate/collect the data. Production transparency should be thought of as a prerequisite for the content of one scholar's data to be truly accessible to other researchers. Analytic transparency is a separate but closely associated concept. Scholars making evidence-based knowledge claims should provide a full account of how they drew their conclusions, clearly mapping the path on the data to the claims.

Now that the discipline is highlighting research transparency as a core ethical obligation for political scientists, the challenge is to help the scholarly community develop incentives and utilities that make research transparency more feasible and rewarding.

The work necessary to follow the guidelines will vary for different quantitative communities. In some areas of quantitative political science, providing such information is already the norm. Among game theorists, for example, formal proofs detail nonobvious relationships between premises and conclusions. Proofs, in this context, are like a computer code that others can use to verify that specific conditions produce specific conclusions. In game-theoretic research communities, nonobvious claims that lack proofs are not considered credible.

In other fields, the documentation and sharing of "do-files" or "code" is less common. Consider, for example, survey research. For survey producers, procedural transparency entails descriptions of case selection, question selection, interviewer selection, interviewer training, and strategies for managing

the same version of the ANES), they have not always had access to how data producers and analysts collected data, created variables, or knowledge of the exact code (i.e., statistical model) that produced published findings. When this material is not available, replication is undermined as is other scholars' ability to evaluate what a quantitative empirical claim actually means. A website called Political Science Replication now collects such instances and, in so doing, reveals many difficulties associated with contemporary replication attempts.[4]

Despite the disappointing record for data sharing in some quantitative communities, promising signs indicate that research transparency is being taken more seriously in important areas of political science. Archiving of procedural materials, research design registries, and revised data citation practices are three ideas that political scientists are pursuing to make research transparency more rewarding and feasible. We describe each of these ideas in turn.

Of the three ideas, archiving of procedural materials is farthest along. Entities such as ICPSR, Dataverse, and the Open Data Project provide means for scholars to share not only their data, but also supplementary materials that allow others to replicate existing findings. Among survey providers, the ANES (production) has provided unprecedented documentation of this kind. For its 2008 studies, the ANES produced dozens of reports on many steps of its data production processes. Its Online Commons provides histories of the evaluative procedures that the ANES used to choose which of more than 3,000 proposed questions to include on its surveys (Aldrich and McGraw 2011). The site also describes many ways in which questions were evaluated including alternate weighting algorithms (DeBell and Krosnick 2009), and how it developed new code frameworks for open-ended responses (Berent, Krosnick, and Lupia 2013).

A second idea that is growing in popularity is requiring research designs to be registered before rather than after such research is conducted (Humphreys et al. 2013). A benefit of registries for researchers is that it allows them to lay claim to

a set of procedures. To see why this is valuable, note that in much empirical research today, only the final version of a multistep analysis is published. This final version is often influenced by well-known publication biases. Because many journals are hesitant to publish null results, scholars tend to send journals only analyses that produce statistically significant findings. Patterns in published articles (Gerber, Green, and Nickerson 2001) suggest that scholars suppress analyses that do not feature significant coefficients. As King (1986) and Lupia (2008) have written, such incentives may lead scholars to engage in "stargazing," the practice of running data through different model specifications until finding a specification that produces statistically significant results on the variables that the scholar wanted to feature. Stargazing is a problem for many reasons, not the least of which is that stargazing can cause the standard errors under-

A complementary endeavor is the Open Science Collaboration's "badge" system (https://osf.io/project/TVyXZ/wiki/home/). This endeavor allows organizations to award scholars "badges" for "open data," "open materials" (e.g., for providing information about case-selection procedures and "do-files"), and "preregistration." Given the increased attention to such matters in recent years, it seems likely that many scholars will want to attach such labels to their work. Such practices can make research transparency more rewarding for individual investigators while also offering credibility benefits to research communities.

CONCLUSION

This article details the value of increased data sharing and research transparency to quantitative social science and fac-

Social scientists who commit to sharing their data and code give broader populations a basis for treating their work as an endeavor that is valuable and worth supporting.

lying common measures of significance to lose the attributes that make statistical significance meaningful (Rubin 2007). So, when scholars show only "significant" results, it can be impossible for readers to determine whether the results have the substantive meaning that the authors claim.

Research registries, by contrast, allow scholars to document practices and findings at many stages of a research agenda. Researchers can post experimental designs, regression models, computer simulation programs, or a list of hypotheses in a registry. Readers of an article or book can then use the registry to determine whether a result is a true characterization of a focal social relationship or whether it is the product of publication-related biases that have no clear theoretical foundation. In other words, they can see whether published designs, models, programs, and hypotheses represent a larger set of inquiries or are cherry-picked because they produce a desired finding. While registries run the risk of embarrassing researchers who are reticent to reveal that they did not derive the best solution to a problem on their first try, such outcomes are a public good. Scholars, particularly those who eventually succeed in discovering important relationships, can help others advance research more quickly by revealing initial and seemingly sensible specifications that turned out to be suboptimal.

The third idea is oriented toward making research transparency more rewarding. Scholars may ask why they should allow others to reap the benefits of research agendas or analytic strategies to which they devoted substantial time and effort. As was the case for scholars who work hard to accumulate data, there are limited professional incentives to share one's "code." To make transparency more rewarding some scholars have proposed revising citation practices. In addition to data citation practices described in the previous section, scholars are also pursuing "code" citation practices. In computer science, for example, many people recognize the value of code. If scholars expected one another to cite their code, there would be greater incentives to make such code available to others.

tors that affect incentives for quantitative researchers to engage in such practices. An increasing population of scholars is recognizing the link between sharing, transparency, and the ability to evaluate scholarly claims. Regardless of the exact process by which change occurs, scholarly decisions to share data and information about the procedures that produced their conclusions are critical to the future of quantitative social science. Social scientists who commit to sharing their data and code give broader populations a basis for treating their work as an endeavor that is valuable, credible, and worth supporting. ∎

NOTES

1. See, for example, Larsen and von Ins (2010) for a comprehensive review of trends in scientific publication and citation broken down by scientific area. See Appendix B-1 of Chiswick, Larsen, and Pieper (2010) for a report on steady growth in the number of social science PhDs granted in the United States and Canada from 1966 to 2006, Ware and Mabe (2009, 5) for statistics on the steady growth of scientific journals and output and (2009, 56) for a chart on journal article use over time.

2. In rare cases an editor may request data during the review process. Data associated with a manuscript under review should be covered by the same expectation of confidentiality that applies to the manuscript itself. If the manuscript is not accepted for publication, the review should not compromise the author's exclusive access to the data.

3. NIH advises: "In preparing and submitting a data-sharing plan during the application process, investigators should avoid developing or relying on consent processes that promise research participants not to share data with other researchers. Such promises should not be made routinely or without adequate justification described in the data-sharing plan" (Office of Extramural Research, National Institutes of Health 2004; see also Inter-university Consortium for Political and Social Research 2012, 13).

4. http://politicalsciencereplication.wordpress.com/. Accessed on September 20, 2013.

REFERENCES

Acord, Sophia Krzys, and Diane Harley. 2012. "Credit, Time, and Personality: The Human Challenges to Sharing Scholarly Work Using Web 2.0." *New Media and Society* 15: 379–97.

Adam, Nabil R., and John C. Worthmann. 1989. "Security Control Methods for Statistical Databases." *Association for Computing Machinery: Computing Surveys* 21 (4): 515–56.

Aldrich, John H., and Kathleen McGraw. 2011. *Improving Public Opinion Surveys: Interdisciplinary Innovation and the American National Election Studies.* Princeton, NJ: Princeton University Press.

Altman, Micah, and Gary King. 2007. "A Proposed Standard for the Scholarly Citation of Quantitative Data." *D-lib Magazine* 13 (3/4). Cited on September 20, 2013 from http://www.dlib.org/dlib/march07/altman/03altman.html.

American Economic Association. 2013. "American Economic Review: Data Availability Policy 2013." Cited June 18, 2013. Available from http://www.aeaweb.org/aer/data.php.

American Political Science Association. 2012. "Proposed Changes to Ethics Guide." Cited September 20, 2013 from http://www.apsanet.org/content_2483.cfm.

Berent, Matthew, Jon A. Krosnick, and Arthur Lupia. 2013. *Coding Open-ended Answers to Office Recognition Questions from the 2008 ANES Time Series Interviews.* Cited September 20, 2013 from http://www.electionstudies.org/studypages/2008prepost/ANES2008TS_CodingProject.htm.

Chiswick, Barry R., Nicholas Larsen, and Paul Pieper. 2010. "The Production of PhDs in the US and Canada." Institute for the Study of Labor (IZA) Discussion Paper 5367. Cited September 20, 2013 from http://ftp.iza.org/dp5367.pdf.

DataCite. 2013. "Why Cite Data?" Cited June 15, 2013 from http://www.datacite.org/whycitedata.

Data Preservation Alliance for the Social Sciences. 2013. "Data Citations 2013." Cited June 15, 2013 from http://www.data-pass.org/citations.html.

DeBell, Matthew, and Jon A. Krosnick. 2009. *Computing Weights for American National Election Study Survey Data.* ANES Technical Report Series, No. nes012427.

Dorch, Bertil. 2012. "On the Citation Advantage of Linking to Data." Cited on September 21, 2013 from http://hprints.org/docs/00/71/47/34/PDF/Dorch_2012a.pdf.

Esterling, Kevin. 2013. "Transparency-Inducing Institutions and Legitimacy." Berkeley Initiative for Transparency in the Social Sciences. Cited on June 7, 2013 from http://cegablog.org/2013/03/20/tss_esterling/.

Fung, Benjamin C. M., Ke Wang, Rui Chen, and Philip S. Yu. 2010. "Privacy-Preserving Data Publishing: A Survey of Recent Developments." *Association for Computing Machinery: Computing Surveys* 42 (4), Article 14 53 pages. DOI=10.1145/1749603.1749605 http://doi.acm.org/10.1145/1749603.1749605.

Gerber, Alan S., Donald P. Green, and David Nickerson. 2001. "Testing for Publication Bias in Political Science." *Political Analysis* 9: 385–92.

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: 1–20.

International Association for Social Science Information Services and Technology. 2013. "Quick Guide to Data Citation." Cited June 15, 2013 from http://iassistdata.org/sites/default/files/iassist-data-citation-quick-guide.pdf.

International Council for Science: Committee on Data for Science and Technology. 2013. "CODATA Data Citation Standards and Practices Task Group." Cited June 15, 2013 from http://www.codata.org/taskgroups/TGdatacitation/.

Inter-university Consortium for Political and Social Research. 2012. "Guide to Social Science Data Preparation and Archiving: Best Practice Throughout the Data Life Cycle." Ann Arbor, MI.

Inter-university Consortium for Political and Social Research. 2013. "Data Citations." Cited June 15, 2013 from http://www.icpsr.umich.edu/icpsrweb/ICPSR/curation/citations.jsp.

King, Gary. 1986. "How Not to Lie With Statistics: Avoiding Common Mistakes in Quantitative Political Science." *American Journal of Political Science* 30: 666–87.

King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research.* Princeton, NJ: Princeton University Press.

Larsen, Peder Olesen, and Markus von Ins. 2010. "The Rate of Growth in Scientific Publication and the Decline in Coverage Provided by Science Citation Index." *Scientometrics* 84: 575–603.

Long, J. Scott. 2009. *The Workflow of Data Analysis Using Stata.* College Station, TX: Stata Press.

Lupia, Arthur. 2008. "Procedural Transparency and the Credibility of Election Surveys." *Electoral Studies* 27: 732–39.

Lupia, Arthur, and Colin Elman. 2010. "Memorandum on Increasing Data Access and Research Transparency (DA-RT)." Submitted to the Council of the American Political Science Association, September.

Miller, Warren. 1962. Letter to Rensis Likert about the Establishment of the Interuniversity Consortium for Political and Social Research. Downloaded on June 6, 2013 from http://www.icpsr.umich.edu/files/ICPSR/fifty/Miller-Likert.pdf.

Mooney, H. 2011. "Citing Data Sources in the Social Sciences: Do Authors Do It?" *Learned Publishing* 24 (2): 99–108.

National Research Council. 2003. *Protecting Participants and Facilitating Social and Behavioral Sciences Research.* Washington, DC: National Academies Press.

National Research Council. 2005. *Expanding Access to Research Data: Reconciling Risks and Opportunities.* Washington, DC: National Academies Press.

Office of Extramural Research. National Institutes of Health. 2004. "Frequently Asked Questions: Data Sharing, February 16, 2004." Cited June 18, 2013 from http://grants.nih.gov/grants/policy/data_sharing/data_sharing_faqs.htm#907.

Pew Research Center. 2013. "Public Split over Impact of NSA Leak, But Most Want Snowden Prosecuted." Cited on September 20, 2013 from http://www.people-press.org/2013/06/17/public-split-over-impact-of-nsa-leak-but-most-want-snowden-prosecuted/.

Piwowar, Heather, and Todd J. Vision. 2013. "Data Reuse and the Open Data Citation Advantage." *PeerJ PrePrints.* Published: 4 Apr 2013, doi: 10.7287/peerj.preprints.1.

Prewitt, Kenneth. 2013. "Is Any Science Safe?" *Science* 340 (6132): 525.

Rubin, Donald B. 1993. "Comment on "Statistical Disclosure Limitation."" *Journal of Official Statistics* 9: 461–68.

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: 20–36.

Ware, Mark, and Michael Mabe. 2009. "The STM Report: An Overview of Scientific and Scholarly Journal Publishing." International Association of Scientific, Technical, and Medical Publishers (Oxford, UK). Cited on September 20, 2013 from http://www.stm-assoc.org/2009_10_13_MWC_STM_Report.pdf.

Xiao, Xiaokui, and Yufei Tao. 2006. "Personalized privacy preservation." In *Proceedings of the ACM SIGMOD Conference.* Association for Computing Machinery, New York.

Zhang, Qing, Nick Koudas, Divesh Srivastava, and Ting Yu. 2007. "Aggregate Query Answering on Anonymized Tables. In *Proceedings of the 23rd IEEE International Conference on Data Engineering.*

Science Deserves Better: The Imperative to Share Complete Replication Files

Allan Dafoe, *Yale University*

In April 2013, a controversy arose when a working paper (Herndon, Ash, and Pollin 2013) claimed to show serious errors in a highly cited and influential economics paper by Carmen Reinhart and Kenneth Rogoff (2010). The Reinhart and Rogoff paper had come to serve as authoritative evidence in elite conversations (Krugman 2013) that high levels of debt, especially above the "90 percent [debt/GDP] threshold" (Reinhart and Rogoff 2010, 577), posed a risk to economic growth. Much of the coverage of this controversy focused on an error that was a "perfect made-for-TV mistake" (Stevenson and Wolfers 2013) involving a simple error in the formula used in their Excel calculations. The real story here, however, is that it took three years for this error and other issues to be discovered because replication files were not publicly available, nor were they provided to scholars when asked. If professional norms or the *American Economic Review* had required that authors publish replication files, this debate would be advanced by three years and discussions about austerity policies would have been based on a more clear-sighted appraisal of the evidence.

An essential characteristic of science is the commitment to transparency. Assumptions should be clearly stated, evidence should be publicly verifiable, and the basis for inferences should be explicit. Independent researchers should be able to reproduce, at least in principle, the structure of inferences linking assumptions, prior theory, other findings, data collection, data processing, and data analysis, to an alleged scientific finding. Much of the institution of science exists to promote transparency, such as the strong norms around citations, the requirement to describe methods, the esteem for formal methods of inference (statistics and formal theory), expectations about maintaining (laboratory or field) notebooks, the expectation to publish proofs of theorems, and the condemnation reserved for nonreproducible results.

Transparency is a foundation for a number of core features of science: *refutability*, *openness*, *cumulation*, and *minimal barriers to entry*. Transparency makes scientific work more *refutable*—more subject to detailed criticism—which is the basis for scientific progress. Transparency makes a scientific enterprise more *open* to exploration by others, facilitating divergent interpretations of results and alternative uses of data and tools. The *cumulation* of data, tools, and findings is essential for the progress of science. The sharing of findings is incentivized through rewards to publications. There are not, however, adequate individual incentives for the sharing of data and tools. The sharing of data and tools is an essential public good for science; a commitment to transparency would promote this public good. Transparency, by making more steps of the scientific process publicly observable, *reduces the barriers to entry* for students and novices. Keeping scientific discussion accessible improves the scientific enterprise by reducing the costs to training new scientists, by bringing in new perspectives, by permitting more cross-disciplinary conversation, and by keeping fields open to external criticism.

This article makes a simple argument. Political science should take its commitment to transparency more seriously by insisting that researchers publish complete replication files, making every step of research as explicit and reproducible as is practical. In return, political science will become more refutable, open, cumulative, and accessible. Science deserves this commitment from us.

This article proceeds as follows. The first section reviews some evidence about the current state of replication practices in political science. The second section elaborates on the benefits of greater transparency through the sharing of complete replication files. Specific recommendations for authors, journals, and universities are provided in the third section. Online Appendix A discusses exceptions to the prior recommendations for confidential, costly, or proprietary data or code. Online Appendix B discusses a proposal for Replication Audits.

The recommendations presented here apply to any domain of science in which some feature of inference could be practically made more explicit and reproducible. These recommendations apply especially to modes of inference that use computers, because any processing involving a computer can be codified and made reproducible. For this reason this article focuses primarily on replication practices in statistical studies, although the recommendations apply equally to computational theory (theoretical models using computer simulations or solutions). Noncomputational modes of inference can also be made more transparent. For example, Moravcsik (2010) offers valuable recommendations to qualitative researchers.

EVIDENCE FROM POLITICAL SCIENCE

What is the current state of replication practices in political science? Gherghina and Katsanidou (2013) found that only 18 of 120 political science journals have a replication policy posted on their websites, to say nothing about enforcing those policies. To provide additional data on the state of replication practices in political science, I collected data on the availability of

replication files for publications in two leading journals—*American Political Science Review* (*APSR*) and *American Journal of Political Science* (*AJPS*)—and scholars' attempts to replicate publications.

Replication Practices at *APSR* and *AJPS*

APSR does not have a policy of requiring replication files, although it encourages them. For example, in the submission guidelines it states that authors "are expected to address the issue of data availability. You must normally indicate both where (online) you will deposit the information that is necessary to reproduce the numerical results and when that information will be posted." By contrast, the policy at *AJPS* as of 2010 states that articles "will not be published unless the first footnote explicitly states where the data used in the study can be obtained," the acceptance letter provides instructions for posting files to *AJPS*'s Dataverse site, and the editor, Rick Wilson, frequently double checks that files are posted and has held up publications that have not posted replication files. The beneficial effects of this policy and editorial involvement are evident in figure 1.

Data was collected[1] on the availability of replication files for recent publications in the two top political science journals, the *APSR* since 2010 and the *AJPS* since 2009. We found that 48% of publications using statistical analysis stated on their first page that replication files were available; we were able to find replication files for 68% of these. We were also able to find replication files for 18% of the publications that did not state that replication files were available.

As figure 1 shows, publications at *APSR* are much less likely than *AJPS* to state that replication files are available, and somewhat less likely to provide replication files. Since 2011, nearly 100% of publications at *AJPS* state that replication files are available, increasing dramatically from 2009 before the new replication policy. This policy seems to have substantially increased the actual availability of

Figure 1

Replication Practices at *APSR* and *AJPS*
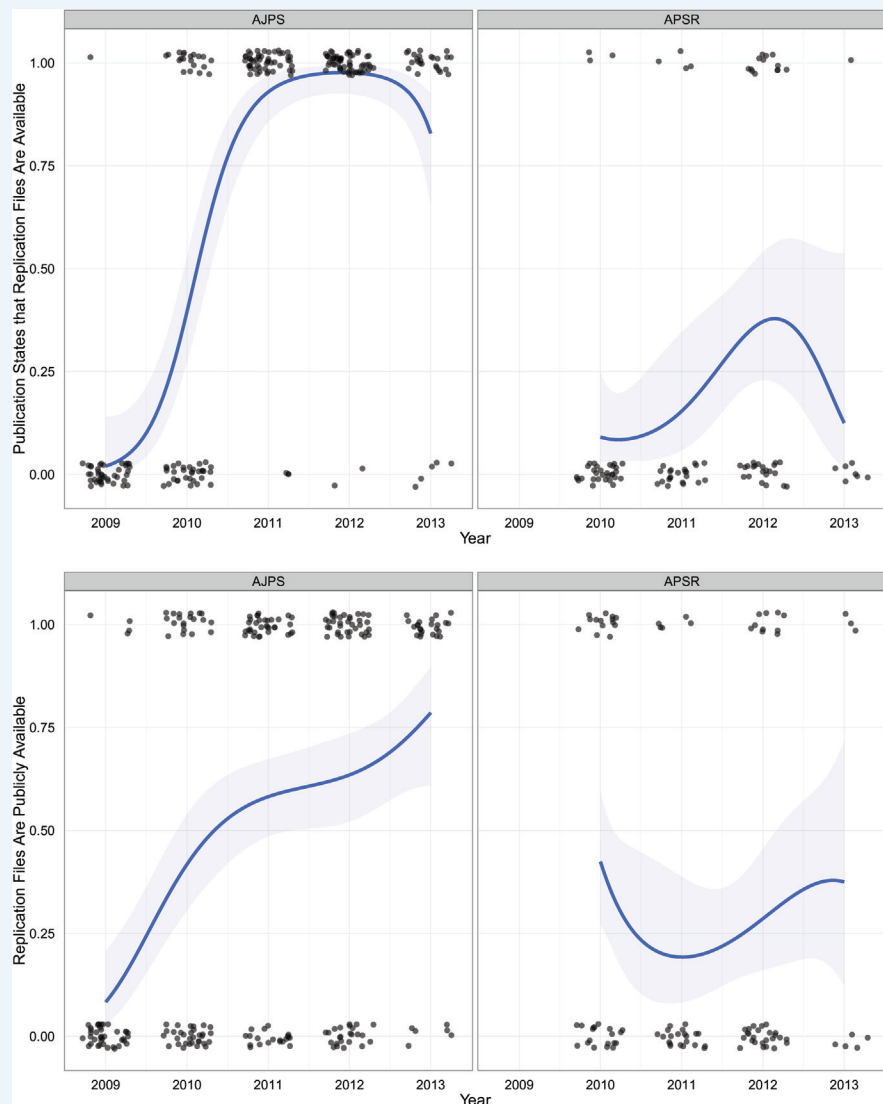
Proportion of published articles employing statistical analysis in *APSR* and *AJPS* that (top row) state that replication files are publicly available, and (bottom row) for which replication files are publicly available. Confidence intervals provide an assessment of whether observed proportions could have come from the same underlying distribution. Results: 30% of the publications at APSR provided replication files, 65% of the publications at *AJPS* 2011–2013 provided replication files; *AJPS*'s 2010 policy substantially increased provision of replication files.

replication files, although we were still unable to find replication files for about 35% of the publications in *AJPS* 2011–2013.

Robustness of Published Results to Replication

The data in figure 1 does not speak to the quality of the replication files that are provided, nor the actual robustness of results. Addressing this question would require a more systematic evaluation of the literature, such as from a Replication Audit (see Online Appendix B).

To offer some preliminary data on this question I surveyed three groups of scholars about their experiences attempting to replicate statistical studies; these groups were students from

my PhD methods class, students from Gary King's PhD methods class, and subscribers to the Political Methodology list-serve. These numbers should be interpreted with caution because it is not from a representative sample: respondents selected into the survey, and respondents selected the work they wanted to replicate. See Online Appendix D for more details about this survey.

This data suggests a mixed conclusion. Of those who responded to the reproducibility of the result, about 52% reported that they were "able to precisely reproduce the main results" and only 13% reported that they were "not able to approximately reproduce the main results." This suggests that many results in political science can be, at least superficially, reproduced, but also that many seem only approximately reproducible. Of those who responded to the robustness of the results,

patterns in psychology and has uncovered incidents of fraudulent data which has led to multiple retractions of articles and the resignation of prominent tenured professors. One psychologist was found guilty of fabricating data for more than 50 publications; this professor did not just "massage" the data, or report only convenient analyses, he literally made the data up and then gave it to his students to analyze for their dissertations (Bhattacharjee 2013). Simonsohn (2012) argues that "requiring authors to post the raw data" will "make fraud much less likely to go undetected."

Fraudulent science probably only makes up a tiny proportion of the scientific output. However, it threatens to dramatically reduce the public credibility of science. Of potentially greater concern (Stevenson and Wolfers 2013) is the unknown proportion of fragile, misleading, or nonreproducible results.

> *The scientific paper is currently a snapshot of a data landscape. Instead, the scientific paper should become an open safari of the data landscape, from which readers are encouraged to depart at any time and explore the landscape on their own. Providing that level of freedom would allow scholars to explore and build on the data to answer their own questions, and otherwise to be much more active participants in the scientific journey of the authors.*

36% reported that "most or all of the key results were robust", 20% that there were "major technical errors though these didn't change the main results", and 56% that results were not robust (responses 5–7). This is encouraging in how many results were found to be robust, while also reinforcing the value of strong transparency norms so that the many fragile results can be more easily uncovered and examined. Also, given that more reproducible and robust work is more likely to share replication files (Piwowar, Day, and Fridsma 2007), these numbers are probably an optimistic appraisal of the reproducibility and robustness of statistical work in political science.

BENEFITS OF SHARING COMPLETE REPLICATION FILES

This section discusses some of the specific benefits of the sharing of complete replication files, which are defined as files that make as explicit and objective as practical every step of research from initial data collection to final statistical output.[2] These benefits include greater refutability, openness, cumulation, and reducing barriers to entry.

A primary benefit of sharing replication files is that it makes research more refutable, and therefore makes the body of non-refuted findings more informative. Fragile, misleading, and nonreplicable[3] statistical analyses can be largely eliminated by the simple requirement that authors be required to submit complete replication files before publication.[4] Doing so will deter many scholars from publishing unreliable analyses, and the scientific community can be relied on to expose many of those remaining.

Consider the cases of serious fraud that have been uncovered in psychology. Uri Simonsohn used data analysis techniques (Shea 2012; Simonsohn 2012) to detect suspicious data

The data reported earlier in this article suggests this proportion is not trivially small in political science.

The sharing of complete replication files has the additional benefit that it opens up scientific research to the questions, insights, and exploration of others. Rather than confine reviewers and readers to a snapshot of the data that has been carefully prepared by the authors, a reader of an empirical analysis with a question or insight could immediately go to the data to evaluate it. Readers might want to double check that primary data has been merged correctly, evaluate how known coding issues have been addressed, identify influential observations, examine particular subsets of the data, implement an alternative conditioning strategy or estimator, or plot the data in potentially insightful ways. Of course, as with any analysis, any findings from such a reanalysis should follow from a principled and well-motivated empirical strategy, and scholars should be wary of the biases from multiple-comparisons (data dredging).

The full set of questions that a reader might have is vast, and cannot, even in principle, be answered in the text of a paper or supplementary materials. However, authors need not anticipate every possible question. Instead, they need only make a sufficient effort to warrant publication, and provide clear and complete replication files so that the scientific community is able to evaluate and build on their work. The scientific paper is currently a snapshot of a data landscape. Instead, the scientific paper should become an open safari of the data landscape, from which readers are encouraged to depart at any time and explore the landscape on their own. Providing that level of freedom would allow scholars to explore and build on the data to answer their own questions, and otherwise to be much more active participants in the scientific journey of the authors.

Scientific productivity exhibits network externalities. The cumulation of data and tools provides substantial benefits beyond the intentions of the creators. A scholar will collect some data for some specific research purpose; however, often those data can be used to answer other questions. This is especially the case for those kinds of observational data, such as cross-national data, that are relevant to multiple research programs. For example, the statistical study of international relations has benefited greatly from the creation of large standardized datasets based on the cumulated work of hundreds of scholars. These externalities are also present in experimental research. Experimental manipulations can be "reused" for down-stream experiments in which other causal effects are investigated.

Sharing complete replication files is also likely to incentivize scholars to be more careful (e.g., see Andrew Gelman here): we face trade-offs in how we invest our time and we are likely to invest more effort in those stages of our research that are

to broadly promote transparency norms such as is articulated in the most recent *Guide to Professional Ethics in Political Science* (APSA 2012): "openness is an indispensable element of credible research and rigorous analysis, and hence essential to both making and demonstrating scientific progress." Specifically, I recommend the following transparency maxim for statistical and computational work.

> Transparency Maxim: Good research involves publishing complete replication files, making every step of research as explicit and reproducible as is practical.

The transparency maxim is both a descriptive statement that good research tends to publish replication files, and a normative statement that the publishing of replication files is a necessary component of good research.

The transparency maxim is likely to be partly self-enforcing. Researchers are more likely to publish complete replication files as they are more technically proficient, more concerned

Transparency Maxim: Good research involves publishing complete replication files, making every step of research as explicit and reproducible as is practical.

most subject to scrutiny. By making more of the research process subject to scrutiny, scholars will have greater incentives to be cautious with those parts of the research process. This incentivizing effect may be one of the primary benefits of stronger replication norms.

Sharing code for analysis and presentation lowers the barrier to entry for students and others, and promotes the dissemination of useful techniques. Students especially benefit from having access to replication code because it allows them to see precisely how prominent scholars execute their empirical analyses and provides an opportunity for junior scholars to contribute to the research frontier (Rich 2013). Instead of needing to "study under" leading scholars to learn their statistical methods, scholars will be able to learn by working through replication files.

Papers that share their replication data and code have greater visibility and more citations (Gleditsch, Metelits, and Strand 2003; Piwowar, Day, and Fridsma 2007). This is probably partly a selection effect, but also probably partly a causal effect. It is much harder to build off of a study for which replication files are not available. In addition, the sharing of replication files provides a public signal about the quality, confidence, and professionalism of a scholar.

Transparent replication practices are a scientific public good: the benefits are large and shared by many, the costs are small but born largely by the authors. While the benefits vastly outweigh the costs, transparency will likely be underprovided unless individual's incentives are aligned with the group's. Strengthening of formal incentives could help, such as if journals, universities, and funders insist that replication files be publicly posted as part of the publication, promotion, and funding process. Ultimately, however, scientific practice follows scientific norms. To incentivize adequate transparency we need

about the quality of their work, more confident in their work, more concerned with the scientific enterprise, and more concerned with being perceived as producing good research. These motives generate a correlation between replication files and good research. For example, Wicherts, Bakker, and Molenaar (2011) report that willingness to share data is positively associated with the strength of the evidence and the quality of the reporting of statistical results. Similarly, some scholars adopt judgmental heuristics based on the availability of replication files (e.g., here).

As this descriptive association becomes stronger, publishing replication files will send a positive signal about the quality of one's research (or the failure to publish replication files will send a negative signal). Low quality research cannot easily "fake" this signal because the very act of publishing replication files makes it much easier to evaluate the quality of the research. This signal will then encourage scholars, journals, and universities that produce good research and wish to be perceived as such to publish replication files. Publishing replication files is thus an informative signal of the quality of one's research.

RECOMMENDATIONS

The following are some specific recommendations about how to produce good replication files for researchers engaged in statistical analysis. This advice is similarly applicable to scholars engaged in computational theory. Some advice is also offered here for journals, universities, and funders about how best to promote these practices. The American Political Science Association has also recently revised its Guide to Professional Ethics, Rights and Freedoms to emphasize and clarify researchers' "ethical obligation to facilitate the evaluation of their evidence-based knowledge claims through data access,

production transparency, and analytic transparency so that their work can be tested or replicated" and offers additional guidance on these topics for quantitative and qualitative researchers. For advice for qualitative researchers, see Moravcsik (2010). Davenport and Moore (2013), Hook et al. (2010), and ICPSR offer advice on data preparation and archiving.

Recommendations for Statistical Studies

1. Do all data preparation and analysis in code. Even if analysis is done by "clicking and pointing," most statistics programs (such as Stata and SPSS) produce the code required to replicate each step.

2. Adopt best practice for coding. Some recommendations to keep in mind are:

   - Use comments and functions to make your code clear. Keep your code clean and clear. Comment liberally to remind yourself and communicate to others what your code is (supposed to be) doing. Use functions to execute specific commands, especially when these commands are repeatedly used.

   - Test your code. Build in routine tests to make sure that your code is doing what you think it is doing. Execute the same procedure in multiple redundant ways to reduce

mary files to the final reported output. In our case, this chain of custody should be fully documented and easily reproducible.

4. Fully describe your variables. Somewhere—in variable labels, a codebook, paper, or comments in the code—the meaning of variables needs to be clearly communicated. The original sources or coding rules for variables should be provided. A reader should be able to trace a variable back to its original creation, and the author who first created a variable should clearly document the rules by which the variable was constructed. It is unacceptable to share data files for which it is unclear precisely where a variable came from, let alone one in which the variables names are an indecipherable character string.

5. Document every empirical claim. Every empirical claim in a paper based on the data should be explicitly produced somewhere in the replication code. This includes all graphics and tables, but also any in-text reference to some feature of the data. It should be easy for others to link empirical claims in the paper to the relevant portion of the code. One way to do this is to include a quote in the replication code of the sentence in the paper in which the empirical claim is made; this way a reader can search for the text of the empir-

Journals are the key site for improving replication practices.

the risks of a mistake. It is not uncommon for results to be driven by a misimplemented routine such as the mishandling of missing data.

   - Run your final code all the way through from scratch. Before finalizing the paper, the entire replication code should be rerun from beginning to end. Make sure you set a seed, and make a log file (in Stata) or use something like knitr (in R), so that this final run is recorded and fully reproducible. Also make sure that all relevant data files are included in the replication files. One way to ensure this is to only call data files from within the folder where the code is stored, and then to upload the entire replication folder for archiving.

   - For a helpful discussion of strategies to improve replication practices and code, see Bowers (2011), Appendix A of Shalizi (2013), and Gandrud (2013).

3. Build all analysis from primary data files. Download data files from the original source, and include a precise reference in your code or paper to this original source. Lock these primary files to prevent accidental changes to them. This way any errors that occur will take place in your code, which can be diagnosed and corrected. Share these original data files along with all other files in your replication files. What good is sharing a final data file and replication code, as many scholars currently do, if the crucial decisions and errors were made earlier in the merging and cleaning of the data? We should take the attitude toward a statistical estimate as legal courts take toward evidence: there should be a clearly documented "chain of custody" from trusted pri-

ical claim in the replication code. Another strategy is to have one's replication code follow the structure of the paper.

6. Archive your files. Upload this finalized set of analysis code and data files to a reliable third-party site such as ICPSR or Dataverse.

7. Encourage coauthors to adopt these standards. Maintaining good replication practices is more difficult when one's collaborators do not have the same replication practices. Scholars may be understandably reluctant to impose strict requirements on their colleagues. I recommend sharing with them papers such as this one that outline good replication practices and the reasons behind them, and otherwise leading by example.

Recommendations for Journals

Journals are the key site for improving replication practices. I describe here a set of policies that journals could readily adopt. This package of policies was crafted to have minimal cost to the journal, to maintain the status quo with respect to when data and code should be shared with the public (at time of publication), and to maximally improve replication practices. Adopting these recommendations will improve the quality of work being published, will signal that the journal has higher standards, and will likely increase the prominence of the journal (Gherghina and Katsanidou 2013).

1. Require complete replication files before acceptance. The simplest policy is a nominal requirement that the authors make replication files available. For example, *AJPS* requires

that "the first footnote explicitly states where the data used in the study can be obtained for purposes of replication." As is suggested by the data in figure 1, this change made a substantial difference for *AJPS* (compare years before and after 2010). However, despite that nearly 100% of *AJPS* articles now explicitly state the location for replication files, less than 75% actually provide replication files. Others have similarly noted the limits of requiring authors to agree to or sign statements of intent to share data (Wicherts et al. 2006; Savage and Vickers 2009). As such, to achieve sufficiently high compliance, journals need to actually ensure that replication files are posted or to post it themselves. For example, the *Quarterly Journal of Political Science* ensures that replication files are available and that it is possible to replicate the results before an article is published. This process of checking or posting replication files can be partly automated in the workflow programs used by journals. Journals may want to follow the journal *Biostatistics* by indicating on the first page of a publication (with a "D", "C", and "R") whether data and/or code is available, and whether the publication has passed a "reproducibility review" (Peng 2011). Journals may also want to adopt the policy at *Nature* of requiring an "accession number" or URL for the replication files at the time of first submission; the files can then be released at the time of acceptance.

2. Encourage high standards for replication files. The journal should articulate its expectations about the quality of replication files to authors. Ideally the journal will encourage high standards, such as those articulated above. Replication files could be made available to reviewers after a revise and resubmit decision, allowing reviewers the option to include the quality of replication files in their assessment of the publication.[5]

3. Implement a replication audit. A replication audit involves assembling a replication team of trusted researchers to evaluate the reproducibility and robustness of a random subset of publications from the journal. By guaranteeing regular space in the journal for the replication audit the journal (1) helps reward the act of evaluating the reproducibility and robustness of published work, (2) incentivizes authors to invest additional effort to make sure that their results are reproducible, robust, and that their inferences are not misleading, and (3) provides a diagnostic of the (hopefully improving) quality of empirical work in the journal. The replication audit is described in Online Appendix B. A replication audit is preferred to the exclusive publishing of replication articles on a case-by-case basis because the latter process is more susceptible to publication bias that will over-represent "interesting" replications that claim to overturn earlier studies.

4. Retract publications with nonreproducible analyses.[6] If an analysis cannot be reproduced, even by the original authors when given ample opportunity to do so, the results from the study cannot be trusted and the study should no longer be a part of the public scientific record. Publications based on nonreproducible analyses should, therefore, be retracted. David Laitin, in a personal communication, recommended the adoption of a standard retraction procedure for politi-cal science. Doing so would make the retraction policy more transparent, remove editorial discretion, and insulate editors from legal retaliation. The standards for such a retraction policy could be set so that only the most egregious cases of nonreproducible analyses are retracted: where the finding cannot be approximately reproduced by the original authors. Even if rarely activated, a retraction policy would promote replication practices by establishing the norm that authors are responsible for providing adequate replication files and by providing strong sanctions against the worst kinds of nonreproducibility.

Recommendations for Universities

1. Universities can provide institutional support for producing complete replication files. For example, the Institution for Social and Policy Studies (ISPS) at Yale University provides a service in which they help produce and publish complete replication files in both R and Stata for ISPS funded research (see here). Harvard's Institute for Quantitative Social Science has built the Dataverse Network Project which provides long-term archiving of replication files and other services.

2. Universities could encourage and expect high replication standards from their students and scholars. Students should be encouraged to submit replication files for course papers. Departments could have a policy recommending publication of complete replication files for all published work.

3. Norms of scholarly evaluation could place more emphasis on transparency and specifically the provision of replication files.

Recommendations for Funders

Require recipients of funding to commit to transparency and specifically to publish replication files. The National Science Foundation, for example, now requires a data management plan as part of any proposal, although of course this is not sufficient (for a satire of one, see here).

CONCLUSION

The study of politics rightly aspires to be scientific: it aims to establish generalizable causal insights from the nonsubjective, replicable, and transparent empirical evaluation of precise and logical theories (Gerring 2011, 11). Relative to many natural sciences, however, political science faces daunting methodological challenges. We are less able to rely exclusively on experiments to resolve many of our questions. We can rarely isolate mechanisms and processes in a controlled setting. Our subjects of study rarely follow simple mathematical patterns, are highly context dependent, and adapt to our interventions and theories. However, political science has better replication practices than many of the natural and social sciences. We should be proud of this and continue to lead the way.

Hammerström, Nicole Janz, Gary King, David Laitin, Daniel Masterson, Gwyneth McClendon, Guadalupe Tuñón, Rick Wilson, Baobao Zhang, Magnus Öberg, and especially the editorial committee at the *Journal of Peace Research*, John Bullock, Don Green, and Arthur Lupia. ▪

NOTES

1. Data was collected by Guadalupe Tuñón, Peter Repucci, and myself.

2. See text in the section Recommendations for Statistical Studies for more discussion of what constitutes "complete replication files."

3. We might distinguish between the "replication of a study" in which the research design is replicated on new data (also called "broad replication"), and the much less informative but nonetheless important "replication of an analysis" (or "narrow replication") in which the analysis is replicated on the same data. Other fields such as computational science distinguish between "reproducibility" (replication of analysis) vs "replication" (replication of a study); I use the term "reproducible" to refer specifically to research for which the analysis is replicable. Sharing of replication files foremost promotes reproducible research (replications of analysis), though it might also promote replications of studies if the greater transparency facilitates the execution of the study on a new sample.

4. This article focuses on replication practices for statistical empirical work. Qualitative scholarship would also be much improved by the adoption of stronger practices of data transparency (see Moravcsik 2010).

5. Another option is to require that reviewers have access to replication files from first submission. However, a number of scholars have expressed concern over this proposal because they are not comfortable with others having access to replication materials so much in advance of publication.

6. I thank David Laitin for raising this idea.

REFERENCES

APSA. 2012. "A Guide to Professional Ethics in Political Science." URL: http://www.apsanet.org/content86135.cfm.

Bhattacharjee, Yudhijit. 2013. "The Mind of a Con Man." *New York Times Magazine*, April 26. URL: http://www.nytimes.com/2013/04/28/magazine/diederik-stapels-audacious-academic-fraud.html?pagewanted=all&_r=1&.

Bowers, Jake. 2011. "Six Steps to a Better Relationship with Your Future Self." *Political Methodologist* 18 (2): 9.

Bueno de Mesquita, Bruce, Nils Petter Gleditsch, Patrick James, Gary King, Claire Metelits, James Lee Ray, Bruce Russett, Håvard Strand, and Brandon Valeriano. 2003. "Symposium on Replication in International Studies Research." *International Studies Perspective* 4 (1): 72–107. URL: http://onlinelibrary.wiley.com/doi/10.1111/1528-3577.04105/full.

Davenport, Christian, and Will H. Moore. 2013. "Conflict Consortium Data Standards & Practices." Manuscript.

Gandrud, C. 2013. *Reproducible Research with R and Rstudio*. Chapman & Hall/CRC.

Gerring, John. 2011. *Social Science Methodology: A Unified Framework*. New York: Cambridge University Press.

Gherghina, Sergiu, and Alexia Katsanidou. 2013. "Data Availability in Political Science Journals." *European Political Science* 12: 333–49.

Gleditsch, Nils Petter, Claire Metelits, and Håvard Strand. 2003. "Posting Your Data: Will You Be Scooped or Will You Be Famous?" *International Studies Perspectives* 4 (1): 9–95. URL: http://gking.harvard.edu/files/replvdc.pdf.

Herndon, Thomas, Michael Ash, and Robert Pollin. 2013. "Does High Public Debt Consistently Stifle Economic Growth? A Critique of Reinhart and Rogoff." Political Economy Research Institute Working Paper Series (322).

Hook, Les A., Suresh K. Santhana Vanna, Tammy W. Beaty, Robert B. Cook, and Bruce E. Wilson. 2010. "Best Practices for Preparing Environmental Data Sets to Share and Archive." URL: daac.ornl.gov/PI/BestPractices-2010.pdf. An earlier version of this was published in: 2001. *Bulletin of the Ecological Society of America* 82 (2): 138–41.

King, Gary. 1995. "Replication, Replication." *PS: Political Science and Politics* 28 (3): 444–52.

Krugman, Paul. 2013. "The Excel Depression." *New York Times*. URL: http://www.nytimes.com/2013/04/19/opinion/krugman-the-excel-depression.html.

Moravcsik, Andrew. 2010. "Active Citation: A Precondition for Replicable Qualitative Research." *PS: Political Science and Politics* 43 (1): 29–35.

Peng, Roger D. 2011. "Reproducible Research in Computational Science." *Science* 334 (6060): 1226–27.

Piwowar, Heather A, Roger S. Day, and Douglas B Fridsma. 2007. "Sharing Detailed Research Data Is Associated with Increased Citation Rate." *PLOSOne* 2 (3): e308.

Reinhart, Carmen M., and Kenneth S. Rogoff. 2010. "Growth in a Time of Debt." *American Economic Review: Papers & Proceedings* 100 (2): 573–78.

Rich, Timothy S. 2013. "Publishing as a Graduate Student: A Quick and (Hopefully) Painless Guide to Establishing Yourself as a Scholar." *PS: Political Science & Politics* 46 (2): 376–79.

Savage, C. J., and A. J. Vickers. 2009. "Empirical study of data sharing by authors publishing in PLoS journals." *PLoS One* 4 (9): e7078.

Shalizi, Cosma Rohilla. 2013. "Advanced Data Analysis from an Elementary Point of View." URL: http://www.stat.cmu.edu/~cshalizi/ADAfaEPoV/.

Shea, Christopher. 2012. "The Data Vigilante." *The Atlantic*, December. URL: http://www.theatlantic.com/magazine/archive/2012/12/the-data-vigilante/309172/.

Simonsohn, Uri. 2012. "Just Post It: The Lesson from Two Cases of Fabricated Data Detected by Statistics Alone." *Psychological Science* 24: 1875–88.

Stevenson, Betsey, and Justin Wolfers. 2013. "Six Ways to Separate Lies from Statistics." Bloomberg. URL: http://www.bloomberg.com/news/2013-05-01/six-ways-to-separate-lies-from-statistics.html.

Wicherts, Jelte M., Denny Borsboom, Judith Katsand, and Dylan Molenaar. 2006. "The Poor Availability of Psychological Research Data for Reanalysis." *American Psychologist* 61 (7): 726–28.

Wicherts, Jelte M., Marjan Bakker, and Dylan Molenaar. 2011. "Willingness to Share Research Data Is Related to the Strength of the Evidence and the Quality of Statistical Results." *PLoS One* 6 (11): e26828.

Research Transparency and Data Archiving for Experiments

Rose McDermott, *Brown University*

Although still more common in medical studies and some other areas of social science such as psychology and behavioral economics, experimental work has become an increasingly important methodology in political science. Experimental work differs from other kinds of research because it systematically administers a specific treatment to part of a population while withholding that manipulation from the rest of a subject pool. The best studies strive to keep all other aspects of the experiment similar, so that any emergent difference between the treatment and control group that emerge provide unparalleled traction in determining causal inference. Many other valuable forms of social research use observation of the natural world, rather than depending on intervention to advance understanding. Because experimentalists can create the environment or process they want to study, this strategy of intervention and manipulation constitutes the main distinction between experimental work and other forms of social observation.

In spite of this critical separation, experimentalists confront some methodological challenges and opportunities that both mimic and diverge from those scholars who are engaged in other forms of qualitative and quantitative work that depends on observation of the natural world. Like those scholars who conduct quantitative research, experimentalists often work with large data sets involving multiple pieces of independent observations, most of which are analyzed statistically to determine results and interpreted in ways familiar to quantitative researchers. However, like qualitative researchers, experimentalists often work with sensitive populations with concerns about protecting individual identities. This work requires that subjects' safety and confidentiality are protected above all other values. In addition, also like qualitative researchers, interviews that take place as part of debriefing may prove informative and useful in ways that require a different kind of data archiving than standard number files provide. Finally, unlike either quantitative or qualitative work, experiments also involve particular experimental protocols, treatment assignments, manipulations, and even Consolidated Standards of Reporting Trials (CONSORT) files that show subject mortality that require unique characterizations. These files differ from other types of work and may require unique standards to achieve research transparency and proper archiving. CONSORT was created to improve the transparency in reporting on randomized clinical trials. The CONSORT statement encourages reporting of a 25-item checklist and flow diagram. This standard reporting strategy provides complete and transparent reporting of all aspects of design, analysis, and interpretation of experimental investigations. The checklist includes title, abstract, introduction, methods, results, conclusion, and supplementary information. The flow chart shows how subjects move through the four stages of clinical trial from enrollment, assignment to condition, follow-up, and analysis.

This article in this symposium outlines some of the current standards and developments designed to achieve increased transparency and archiving of experimental work. The following section discusses some areas of consensus in reaching this goal as well as some challenges that confront scholars who wish to pursue experimental work, particularly in the context of field experiments. The second section outlines some potential strategies and next steps that may be useful to maximize research transparency and data archiving in concert with the goals pursued by other research traditions in political science.

RECENT DEVELOPMENTS

This brief overview examines recent attempts to increase transparency and archiving in experimental work, efforts particularly evident in at least two distinct areas. First, the new *Journal of Experimental Political Science (JEPS)*, initiated under the auspices of the APSA's Experimental Methods Section, whose initial editors will be Rebecca Morton and Joshua Tucker at New York University, has established clear standards for submission, review, and the conduct of research. Second, similar specific standards have been pursued and endorsed by Experiments in Governance and Politics (EGAP), which has tasked itself with "supporting experimental research on the political economy of development." This project encompasses many prominent scholars working in both the laboratory and field experimental areas and has been spearheaded by Jeremy Weinstein at Stanford University and Macartan Humphreys at Columbia University, among others. Although many other efforts exist, these two have been the most systematic attempts to assemble scholars working in experimental traditions and secure commitment to follow particular procedures designed, in part, to achieve transparency, accountability, and replicability.

A third effort has recently emerged: the Berkeley Initiative for Transparency in the Social Sciences (BITSS). Although started by a group of scholars primarily interested in development studies, BITSS encourages research transparency across a wide array of social science disciplines, including political science. In particular, it promotes study registries, data sharing, and replication through learning, discussing, and disseminating best practices. While many members of its leadership are experimentalists, BITSS does not limit itself

to experimental processes and procedures, nor does it restrict the content to development work as EGAP largely does. Many of its suggestions and strategies are applicable to advancing research transparency in experimental work, but BITSS's dictums supporting transparency and replication extend beyond interventional research into observational methods as well.

The new *JEPS* established a set of instructions for both contributors and reviewers that directly speak to many of the issues raised in this symposium. In addition, it directly endorses the reporting standards for experimental work developed by the Experimental Methods Section's Standards Committee. The Standards Committee that developed this document was headed by Alan Gerber and included Kevin Arceneux, Cheryl Boudreau, Conor Dowling, Sunshine Hillygus, and Tom Palfrey. These standards addressed many aspects of experimental design that affect experimental treatment across both laboratory and field settings. Specifically, these standards ask authors, first, to clearly state their hypotheses. Next, authors are asked to explicitly state which subjects were included or excluded from consideration, how and where they were recruited, and to provide the dates when the study was conducted. If a survey was used, authors are asked to supply response rate. Although not stated, best practices would expect that the survey instrument should also be provided, even if only as part of an online supplementary index. This subject information is crucial for achieving appropriate levels of research transparency, more so in experimental work because experimental work proceeds largely through aggregation that occurs as both subject population and context are expanded or shifted. Proper replicability procedures demand that subsequent researchers are aware of the previous populations that have been investigated, and the context under which they have been examined.

Then, the standards request that authors provide statements of what are called their "allocation method." This refers to information regarding whether and how processes of randomization were used in the experiment; experimenters are asked to provide information about how this was accomplished and evidence that it was achieved. Scholars are also asked to provide information about whether subjects, administrators, and analysts were blind to the conditions of the subjects across treatments. In addition to standard requirements about detailing the conditions of treatment and control, providing the instruments of measurement and assessment, following careful standards of analysis, noting institutional review board (IRB) approval and whether or not deception was used, the most unusual and potentially controversial standard asks authors to provide a CONSORT flow chart detailing how many subjects were lost across the course of the study by treatment condition. Where low noncompliance exists, authors are instructed that they can omit the diagram and replace it with a statement in the text. The goal here is to make clear how and why certain subjects may have dropped out of one condition more than another, possibly indicating a systematic difference in who is affected by the treatment and why that might otherwise be lost if only completed subjects were analyzed and presented in the final results. Although this requirement is often considered in medical experiments—where, for exam-

ple, drug side effects may cause more patients in one condition than another to drop out of a study and this information may be crucial for issues of patient compliance—this standard is not typical in either the psychological or economics experimental literature. The political science standards discussed here go one-step further to require authors to report statistics for intent-to-treat, which is another technique for ensuring that results do not reflect biased findings by failing to incorporate those lost to analysis at earlier phases of the experiment.

In its instructions to contributors and reviewers, *JEPS* goes beyond the standards for experimental work provided by the Standards Committee. Because the journal requests shorter articles, *JEPS* notes that some of the material required by the standards may be uploaded into online supplementary material so that it does not count against manuscript word count but will still allow other scholars to find the information. In addition, they require not only evidence of IRB authorization but also disclosure of potential conflict of interest; of course, this is very important in any cases where scholars also have a financial interest in companies that run surveys or experiments for profit. Perhaps most innovatively, a review history of the manuscript, which details which journals the article has previously been submitted to and their responses to requests for revision, is required. This may allow work, which was rejected for lack of wider audience interest or lack of experimental sophistication on the part of reviewers or editors, to receive more expedited review. Finally, the journal does require, for replication purposes, that all data relevant to an experiment be submitted. The instructions proceed as follows:

> For experiments these files should include original experimental instructions or other experimental instruments used in the experiments such as surveys, videos, computer programs, etc., and the raw data from the experiment. For empirical papers, both using experimental or observational data, the final data set(s) and programs used to run the final models, plus a description of how previous intermediate data sets and programs were used to create the final data set(s) must be provided … Authors must provide a Readme PDF file listing all included files and documenting the purpose and format of each file provided, as well as instructing a user on how replication can be conducted. If a request for an exemption based on proprietary data is made, authors should inform the editors if the data can be accessed or obtained in some other way by independent researchers for purposes of replication. Authors are also asked to provide information on how the proprietary data can be obtained by others in their Readme PDF file. A copy of the programs used to create the final results is still required.

Similar instructions are provided to reviewers, noting most unusually that the journal encourages the submission of replication studies and null findings, thus explicitly encouraging authors to submit work designed to replicate studies conducted by other investigators, or studies whose lack of findings can save others from wasting time undertaking work others have already found to be unsubstantiated.

An additional major effort designed to achieve consistency, replicability, transparency, and accountability in experimental work has been undertaken by those scholars involved

in EGAP. Most of these studies, because of the content they examine, tend to take place in field contexts. In addition, although the explicit subject goals of this group are designed to investigate issues specifically related to political economy, governance, and development, many of the methodological issues they confront and address do not differ substantively from those facing any experimentalist.

Two main aspects are central to their campaign to enhance transparency and archiving. First, like *JEPS*, EGAP has developed a set of standards adopted by the unanimous vote of the membership. Endorsing these standards is a condition of membership for incoming new members. This statement is relatively simple and straightforward and encompasses human subject protection, transparency, rights surrounding review, and publications of data and findings and remuneration, which, while discouraged at the very least, must be disclosed. This last item appears similar to the conflict of interest statement requested by the editors of *JEPS*.

The second aspect revolves around various strategies designed to institutionalize procedures to ensure transpar-

enter into the overall literature, but many scholars may continue to reinvent the wheel, not knowing that previous work has shown that certain speculated relationships fail to exist. Another kind of registration system, which would require full-scale review of an overall research design, might require journal editors to make publication decisions based on the design prior to the collection of data so as to prevent such bias at the back end. Under this scheme, when accepted, journals would be required to publish papers regardless of the outcome of results. It is easy to see why journals may not want to comply with a strategy that ties its hands in this way, because there may be other reasons, including something as simple as bad writing, which may incline the the journal to eschew publication at the final stages for reasons not evident in the presentation of the design.

The public nature of these prior commitments would severely reduce the incentive, and dramatically escalate the humiliation, associated with violating the original research goals set out by the experimentalists. One can imagine other kinds of registration strategies designed to "name and shame"

Similar instructions are provided to reviewers, noting most unusually that the journal encourages the submission of replication studies and null findings, thus explicitly encouraging authors to submit work designed to replicate studies conducted by other investigators, or studies whose lack of findings can save others from wasting time undertaking work others have already found to be unsubstantiated.

ency. Many of these involve various kinds of registration opportunities. For example, scholars are encouraged to register their preanalysis plans to reduce the likelihood of "data fishing" or what in the old days of social psychology used to be called "dust bowl empiricism," which can become a serious problem especially with the tools available with current computing power. This registration tool, which already has been substantially used, allows scholars to state which aspects of their data they will analyze in which way, detailing both hypotheses and methods of analysis; these records are then freely available on the EGAP website for other scholars to see.

The typical posting on the EGAP website lists the particular hypotheses that experimentalists plan to test. However, a continuum of registration demands or designs could be included. Simple registration seeks to prevent the problem of scholars cherry-picking those aspects of their data that show the best or strongest results, or confirm a particular theoretical or ideological position. This registration may help keep authors honest about what they plan to investigate, although it may also unnecessarily restrict creativity by preventing the credible examination of true surprises that can emerge in the context of any data collection. Moreover, this simple registration strategy does nothing to address the problem of publication bias among journals that remain stubbornly resistant to publishing null results in particular. Without adequate representation of the full range of outcomes, not only does bias

norm violators providing successful avenues by which scholars can subtly, but powerfully, strengthen best practice norms in experimental methods.

One of the most interesting additional initiatives is the Transparency and Accountability Initiative funded by a host of high-profile private organizations as well as nongovernmental organizations, including the Ford Foundation, Open Societies, and the Hewlett Foundation. Although focused primarily on achieving these goals in the area of international development, this initiative appears designed to provide mechanisms which can allow citizens to hold their governments accountable through a wide variety of educational, technological and policy innovations. The link to this initiative can be found here: http://www.transparency-initiative.org/about.

CONVERGENCE AND CONTENTION IN BEST PRACTICES

Points and patterns of consistency appear to be clearly emerging in experimental research designed to enhance research transparency and data archiving. First, most notably, experimentalist largely consider both of these things to be not only good things, but necessary for their own research to proceed apace. Specifically, most experimentalists, perhaps more than those working in other research traditions, know that experimental work proceeds through a process of aggregation and replication, whereby findings from previous work are extended to new populations or within different contexts. For this work to be done well, it must be done carefully, to determine the

limits of particular phenomena and to understand the nature of particular contingencies on expected results. In other words, endorsing and enhancing these practices within the community of experimentalists improves everybody's work, and efforts that reinforce individual incentives are often easiest to encourage and expand.

Aside from issues related to transparency and archiving, experimentalists also seem to strongly endorse issues related to achieve accountability. This is most notable in the items related to requiring IRB approval for human subjects, but also in the statements revolving around conflict of interest.

Although political science did not traditionally require that data sets be mounted with publication, scholars who wanted to replicate studies could typically request such data from the authors, and authors might note that such information was available on request in a publication. But as standards across disciplines converged toward posting data with publication, political science journals are increasingly moving in that direction as well. Innovations including advance registration, such as that offered by EGAP, provides even higher standards to which scholars can hold themselves accountable even prior to analysis, write-up, or publication.

Second, experimentalists across the board, whether based in the lab or field settings, clearly endorse the protection of human subjects. This extends beyond the cynical enlightened self-interest that recognizes that abused subjects talk to others and can make future experiments more difficult at best, and rain down lawsuits at worst. However, even in places where IRB approval is not yet the norm, such as many institutions in Europe, scholars recognize that well-treated subjects are not only more cooperative but also supply more accurate information, not only in their experimental responsive, but in the often crucial insights they can provide in proper debriefing procedures.

This human rights issue, however, does raise concerns related to subject confidentiality. Even when every reasonable effort is made to protect subjects' identity, the consequences of exposure may feel great to some subjects, particularly when studies are conducted in war-torn or contentious regions, or across conflictual groups, as often occurs in examinations of interethnic discrimination or civil war. When subjects feel that exposure can be easily gleaned from the sensitive nature of the questions or the idiosyncratic nature of truthful responses, subjects may be understandably reluctant to participate, or to give accurate responses. More important, investigators who include such people really may be placing them at risk, and thus the obligation to protect under such circumstances becomes particularly acute. Investigators who are genuinely concerned about negative consequences devolving to any of their subjects should not include such individuals in their studies, even if significant costs redound to the study. Exclusion under such conditions remains the only ethical path. However, determining when such conditions may arise or be in place may not always remain obvious and the subject's perception must always take precedence over the judgment of the investigator.

I learned a searing lesson in the perception of identity that has stayed with me ever since when I conducted my war games at Harvard. I was taking a variety of measures, including saliva for hormonal analysis, and a copy of their handprint to mea-

sure finger length ratio. I wanted subjects to have an id number that was not their name but that they would remember over several months because of the panel nature of the study. So I used the standard used in VA studies that involved the last four digits of a person's social security number, which are uncommon enough to make replication in a small set rare, but not so unique as to be identifying. On the second day of the study, a young African American woman came in and I started to explain the protocol to her and she physically pulled back and said, "Wait. You want my DNA, my fingerprints, my Social Security number for a study funded by the Department of Defense and you're telling me this is anonymous and confidential? And why am I supposed to believe you?" I was stunned, but I instantly saw how the experience looked completely different than my intent when seen through her eyes. More for my sake than hers, I asked her, without requesting any data, what I could do that would make her feel comfortable. She said she was not sure. I asked if she would feel better if she could pick her own id number. She nodded. She picked a number I still remember for its simplicity a decade later but the point was not that it could not be guessed; the point was that she picked it, not me. I then explained about the copy of the hand and she looked at me and said, so, if you blacked out my fingerprints, you would still want it?" I said yes. She copied her hand. I blacked it out, but I went one step further. I measured what I needed in front of her and took the number and then destroyed the copy while she watched. The information did not change, and the DNA I could extract from the saliva (but did not) was what it is: a totally unique identifier that could never be anything less because of its nature. But I had a completely different understanding of the nature of subject identity and the sensitivity and responsibility involved in protecting individuals not from what I would do, but from what they feared I could do.

However, issues related to protecting the identity of experimental subjects does remain distinct from the graver risks that may accompany the kind of in-depth interview work typically conducted by qualitative researchers. In experimental cases, the easiest way around subject identification, baring the use of biological data, is to never collect subject names; simply assign id numbers that tie relevant linked data together. Anyone who may want to know the identity of participants will never be able to ascertain this information because it was never collected. This becomes an issue, for example, when universities tried to use such information to find students who were in violation of immigration laws to pursue orders of deportment against them. If names that link status to a particular individual are never taken, such protection is ensured even if suspicions arise. With qualitative researchers using interview data, information may reveal the identity of a subject even baring the collection of a name because of the specificity of the information provided; this poses greater risks for the participants and greater challenges for the researcher. This topic is dealt with more in the Colin Elman and Diana Kapiszewski contribution on qualitative research.

This dictum to never collect subject names may run afoul of traditions in both survey research and economics that require compensation of subjects in a way that requires subject

identification. The typical way that this is addressed is to keep two distinct logs that are not merged, one which contains the data, and the second which contains the names of subjects along with contact information for purposes of remuneration or reimbursement. However, constructing a wall between names and data is not always feasible or even successful, often for reasons as simple as the order of entry between the two being matched one to one by research assistants who may not be familiar with the importance of confidentiality. The more unassailable way to address these concerns is to remunerate subjects on sight, either with cash or gift cards chosen from a menu of options. If this recompense needs to be provided through the Internet, it can be done through the generation of randomly assigned codes that can then be redeemed for particular rewards or benefits.

The unsettled challenges that remain at the level of large-scale norms seem to relate to the proprietary nature of data, which is an issue not exclusive to experimentalists. First, any scholar who expends tremendous time and effort designing and conducting a study and collecting data may not want to give it all away before they have had a chance to fully explore all of their potential findings. In this case, by parsing data into pieces and publishing and posting findings from particular pieces and parts of data this can be partially avoided. This strategy is not always viable, especially if many parts of the data are linked theoretically or empirically. Under these conditions, incentives can pull in opposite directions when scholars want to publish early but also to protect their data. In such cases, the researcher may have to decide how to approach these constraints on a case by case basis. Moreover, such a strategy can also run contrary to the expectations or demands imposed by the various registration strategies discussed previously. Although the timing of data release remains distinct from actual research design strategy, and thus need not be delineated in advance in a registry, scholars need to think seriously during the design phase of their work about how they might need or want to parse up their findings during the write-up and publication phase. While it has become increasingly common to break studies into ever smaller parts in search of the ever-larger quantities of publication demanded by promotion and tenure committees, work should be divided according to its conceptual or theoretical specifications rather than its strategic value.

The second case is more common and relates to the embargo standards imposed by Time Sharing Experiments in Political Science (TESS) funded by the National Science Foundation (NSF). When scholars have their proposals accepted and run by TESS, they get the data first, but after a year of embargo, it becomes publically available as mandated by all taxpayer funded work. This means that if investigators do not complete their work in this time, other scholars seeking data can use it. Although often such data goes unused by both investigator and observers, as Diana Mutz's book *Population Based Survey Experiments* (2011) illustrates so well, the time-limited nature of the data embargo does pose a risk to experimentalists who may lose the ability to publish first on their own data.

FUTURE CHALLENGES

The challenges that seem to confront experimentalists pursuing best practices and high standards for research transparency, accountability, replicability, and data archiving overlap with the challenges facing both qualitatively and quantitatively oriented scholars. And here overlap appears to be the key word. One of the risks with various groups pursuing the same agenda in different ways is that the norms that develop may become haphazard or too narrow in orientation. Specifically, for norms to become widely accepted, they must have wide adherence, and when various groups each develop standards and practices independent of one another, but seek to impose their particular branding on their contributors, reviewers, or participants, regulation may become burdensome rather than protective, especially if such rules and procedures have significant areas of disagreement or neglect. Such territoriality may work in opposition to the larger goals, as we have learned often happens in domestic and international politics as well.

Only the most disciplined scholars can achieve true freedom. Creativity does not result from luck or serendipity. Rather, it emerges when a prepared mind encounters unexpected processes in the midst of recognized patterns and structures. Just as it takes a dancer years and years to develop the physical prowess, muscle strength, and skill to express truly original movement, it requires the most tedious discipline and practice as a scientist to develop the experience and talent required to know when deviations from the standard will lead to total failure and when it just might instigate the spark of discovery known as genius.

Only the most disciplined scholars can achieve true freedom. Creativity does not result from luck or serendipity. Rather, it emerges when a prepared mind encounters unexpected processes in the midst of recognized patterns and structures. When best practices become habit, time and energy need no longer be spent on organization and logistics but rather can be allocated to the recognition or generation of such patterns and dynamic processes. Just as it takes a dancer years and years to develop the physical prowess, muscle strength, and skill to express truly original movement, it requires the most tedious discipline and practice as a scientist to develop the experience and talent required to know when deviations from the standard will lead to total failure and when it just might instigate the spark of discovery known as genius. Best practices and norms of transparency and accountability may need to be tailored to specific sub types of particular research methodologies. However, the broader goals need to be shared by journal and press editors, organized sections, and the wider political science community if they are to be adopted as functional and effective norms. Achieving consistency may be a campaign beset by obstacles, but accomplishing the successful adoption of widespread norms of research transparency, data archiving, accountability, and replicability is a goal worth striving for because it not only serves us as academics, helping us conduct better work and receive more credibility from the larger research community, but it also should allow us to communicate our results with more confidence, accessibility, and assurance to our students and the larger public. ■

Making DA-RT a Reality

Thomas M. Carsey, *The University of North Carolina at Chapel Hill*

Calls for greater data access and research transparency have emerged on many fronts within professional social science. For example, the American Political Science Association (APSA) recently adopted new guidelines for data access and research transparency. APSA has also appointed the Data Access and Research Transparency (DA-RT) ad hoc committee to continue exploring these issues. DA-RT sponsored this symposium. In addition, funding agencies like the National Institutes for Health (NIH) and the National Science Foundation (NSF) have expanded requirements for data management and data distribution. These pressures present challenges to researchers, but they also present opportunities.

I define data access as the degree to which scholars who produce a research product (such as a published paper) make the data used in producing that research product available to others. Such access might be limited to just the subset of data used to produce the research product in question, or it might include the entire data set. Access might require signing a data use agreement, collaborating with the original research team who collected the data or waiting until an embargo period has cleared. Proprietary, privacy, or other issues may also limit or prevent data access. In short, there are many shades of gray in evaluating data access.

I define research transparency as the degree to which the process used by scholars to produce a research product is made clear and open to others. Data access is part of research transparency, but it also includes clear descriptions of and access to codebooks, decision rules for collecting, coding, and analyzing data, and, where appropriate, computer code. To borrow a metaphor, research transparency involves describing and providing access to both the ingredients used in producing a research product and the recipe for combining them.

Fostering greater data access and research transparency rests on a strong normative foundation. It increases the exchange of ideas, expands learning from individual studies, permits greater public scrutiny of results, and expands the impact of research in both academic and nonacademic settings. At a more fundamental level, the ultimate goal of research is to generate new knowledge and disseminate it widely. Scholars search for knowledge, but they must share their discoveries, subject them to the scholarly judgment of others, and permit others to build on them. Knowledge advances collectively, not individually, and this collective effort requires data access and research transparency. These principles increase the credibility of academic research and provide greater legitimacy to the research process. Ultimately, adherence to data access and research transparency principles increases the value of the research we produce.

The articles making up this symposium offer insights on a range of issues associated with data access and research transparency. In this article, I turn the conversation toward concrete actions individual scholars and scholarly organizations can take if they accept the premise that data access and research transparency are essential to the collective production of knowledge.

Scholars have limited time and resources, and they face a broad and growing set of demands, constraints, and pressures from a host of sources. In simple terms, scholars cannot do everything they might like, or that we might like of them—they must make choices that involve trade-offs of time, resources, and effort. Universities, academic professional associations, journal editors, and publishers also face time and resource constraints. As a result, concrete proposals to promote greater data access and research transparency must affect the cost-benefit analysis scholars, editors, publishers, and universities use when making choices about how to allocate their efforts. Although I noted earlier a strong normative rationale for greater data access and research transparency, in this article I present some ideas I hope will help to increase the benefits to scholars that come from providing greater data access and research transparency while lowering the costs of doing so.

Finally, while data access and research transparency touch all aspects of social science research, the articles in this symposium highlight important differences between traditional quantitative digital data, qualitative nondigital data, and the various mixes in between. Most of the issues I discuss here are relevant for all kinds of data, although their direct application might be easiest to envision in relation to research based on the application of some data reduction algorithm or similar procedure to digitized information for revealing particular patterns or attributes in the data.

EXPANDED VISIBILITY AND IMPACT

Many scholars advocate greater data access and research transparency because they want to promote research that replicates one or more existing studies. For some researchers, this raises the specter of facing public critique or refutation of their own research claims. From that perspective, some scholars might view providing greater access to their research data as a potential risk. This fear must be countered by recognizing that research that fosters replication is, by definition, influencing the larger research community. Even work that is eventually refuted provides at least part of the impetus for the subsequent critique. Common folklore asserts that the modal number of citations a refereed publication receives is zero. If that is anywhere close to reality, we should place value on research that stimulates replication studies even if its main conclusions are subsequently refuted. Making existing studies easier to replicate also makes them easier to build on, increasing the impact of the original study. Remember, knowledge advances collectively.[1]

If replication is critical to the progress of knowledge, scholarly journals should be encouraged to publish replication studies, and departments and universities should give credit to researchers who produce such work. Doing this would encourage more scholars to engage in such activities. What constitutes a publishable replication study is an open question.[2] Space in leading journals may not be best used for replication studies that use the exact same data and exact same methods to reproduce the exact same results as a previously published paper. However, scholars should be encouraged to confirm or challenge findings using similar data and similar methods as a means of assessing the robustness of a published result. Efforts that uncover findings that cannot be replicated also deserve publication. Foreshadowing a point raised later in this article, the publication of replication papers might be effectively accomplished through publishing such materials online.

We need more research to document the impact of promoting data access and research transparency principles. Some evidence indicates that articles that provide easy access to rep-

We socialize graduate students and junior faculty to respond to these metrics, and successful ones do. If the same value were attached to the production and distribution of data sets, scholars would face the same incentives to produce and share data sets that they currently face to produce and publish papers.

A recent development at NSF magnifies this point. Grant submissions to NSF have long required short biographical sketches for the principal investigator (PI) and all co-PIs that are meant to help reviewers evaluate the qualifications of a research team. In 2012, NSF changed one of the required biographical sketch subsection headings from "Publications" to "Products," explicitly identifying data sets as one of the types of research products scholars might list. Such products, including data sets, must be citable. This decision by NSF acknowledges the value of research data in its own right, but it also highlights the need to provide scholars with a method of data citation. Several efforts are underway to provide rules and methods for data citation—I am most familiar with the effort connected with the Dataverse Network (DVN) project.[3] A

If replication is critical to the progress of knowledge, scholarly journals should be encouraged to publish replication studies, and departments and universities should give credit to researchers who produce such work.

lication data and code are more influential than comparable articles that do not (Piwowar, Day, and Fridsma 2007). Similarly, the use and reuse of data sets enhances the visibility of the original project (Pienta, Alter, and Lyle 2010). However, we do not know if the adoption of more visible data access and research transparency policies increase the visibility, attractiveness, and impact of the journals that do so. We also need research that explores the metrics for measuring the impact of data access and research transparency policies. Universities are making greater use of citation counts to evaluate scholarly impact, and services like Google Scholar make gathering such information easier. If we want to encourage greater data sharing, citations to data sets should be part of a scholar's citation count. This also raises the question of whether all citations should count equally. If scholars cite Article A in a string of five or six citations to support an important, but secondary, point, but use Data Set B in their central analysis, one could argue that Data Set B has a greater scholarly impact than does Article A. Including citations to data sets in a scholar's citation count, and developing a metric to assess the impact of a data set shared by a scholar, would promote greater sharing of data.

DATA AS A RESEARCH PRODUCT

Hiring, promotion, and tenure committees at universities evaluate the actual and potential research productivity of scholars primarily in terms of the papers, books, and other publications they produce. Some also give weight to grants sought or received. The weight placed on each publication or grant is often affected by its perceived actual or potential influence.

byproduct of these efforts, should they succeed, will be metrics for evaluating the impact of data production and data sharing.

Getting data sharing and data citation counts included in hiring, promotion, and tenure decisions will require leadership. Professional associations, like APSA, could make public statements regarding the value of doing so. Leading universities/departments might help establish a trend by adopting such strategies and encouraging other universities/departments to follow. Internal and external reviewers could highlight data contributions made by candidates for promotion and/or tenure in their evaluation letters. I do not support making a significant change in expectations for those already approaching tenure and promotion decisions, but this is one area where strong incentives to promote greater data access and research transparency could be established.

LINKING ARTICLE PUBLICATION WITH DATA PUBLICATION

Sharing data openly as a public good differs from publishing data as a product of research. The publishing industry is grappling with these issues via debates about open access to journals, open publishing, and the like. As the open access debate unfolds, it provides a good opportunity to consider linking the distribution of research findings and research data.

We have well-established workflows for authors, editors, reviewers, and publishers for the publication of academic articles.[4] These workflows change as technologies change, and some important differences exist across disciplines and journals, but for the most part, these workflows are well understood by nearly

everyone involved. First-time authors and first-time editors face a learning curve, but publishers, former editors, and experienced colleagues are readily available for consultation and support. In short, we know how to publish papers.

We know much less about how to publish data sets. For example, Gherghina and Katsanidou (2013) report that only 19 out of 120 journals in political science and international relations have a published policy on data accessibility. Several efforts are under way to help foster better integration between paper publication and data publication, but a great deal of work remains to be done.[5] The goal is to make it easier for authors, editors, and publishers to publish data linked with research articles that use that data. Doing this requires addressing both technical and workflow issues.

Regarding workflow issues, several questions require answers. For example, should authors be expected to submit replication data and code as part of their initial submission, only when invited to revise and resubmit, or only when a paper is accepted for publication? Similarly, should reviewers be asked to review the data and code as part of evaluating a paper under consideration for publication? How much access to data should reviewers be granted prior to publication? Does

issues, and organizations, like the Odum Institute, are involved in some of these efforts. Asking authors, editors, reviewers, and publishers to deal with replication materials necessarily increases their workloads. For these efforts to succeed, we must produce workflows and related tools that make this work as easy as possible for all involved.

IMPLICATIONS FOR RESEARCH TRAINING

Lasting adoption of data access and research transparency principles requires that we integrate these values into our graduate training programs. To do it well, this training needs to start in the standard scope and methods course that most graduate programs offer in the first semester. Such courses often consider research ethics, competing notions of science, and various methods of doing research and collecting qualitative and quantitative data. These courses should incorporate the values of data access and research transparency from ethical and scientific perspectives, but they should also explore developing the associated pragmatic skills. The more that ideas associated with data access and research transparency are blended with discussions of developing research questions, formulating initial research plans, and developing research designs,

Authors, editors, and publishers need a single interface that integrates article submission and publication with data submission and publication. This would make it easier for journals to adopt and monitor replication policies, easier for authors to comply with those policies, and easier for other researchers to find published data and articles related to their own work.

access to data reduce anonymity of authors for journals using double-blind review? Should journal editors or publishers be expected to verify replication materials? What happens to replication materials if a submission is ultimately rejected for publication? Do editorial and production staff have the necessary expertise to evaluate and manage the review of replication materials and data publication?

On the technical side, the development of online article submission platforms for peer-reviewed journals has been a huge benefit for both authors and editors.[6] Similarly, tools like the DVN provide individual researchers with access to a web-based submission system for archiving and sharing research data. Both the DVN and commonly used journal submission systems allow for posting supplementary documents, which means, in theory, that journal submission systems could accept replication materials and the DVN could accept reprints of published papers.[7] However, no platform currently exists that integrates the features of both types of systems. Authors, editors, and publishers need a single interface that integrates article submission and publication with data submission and publication. This would make it easier for journals to adopt and monitor replication policies, easier for authors to comply with those policies, and easier for other researchers to find published data and articles related to their own work. Again, development projects are underway to resolve some of these

the easier it will be for students to incorporate these principles in their work.

Beyond this initial course, most PhD programs in political science offer one or more quantitative methods courses, and many offer additional courses in both qualitative and quantitative methods. These courses vary in their focus on methodological theory versus application, but they often devote little or no time to broader issues of data management, data access, and the generation of transparent research replication materials. Whether the task involves proper documentation of the R code used to estimate a statistical model or proper documentation of field notes from a participant-observation study, these types of applied skills need to be folded into our methods training.

One growing trend within quantitative methods sequences is assigning students the task of replicating an existing published study, with sometimes the added element of providing some extension to that study as well. I have given such assignments regularly for the last decade. Unfortunately, one of the lessons students generally learn from this is how poorly existing research is documented and how difficult it is to replicate published results. Data is not made available, different versions of the data exist but are not clearly documented, decisions used to transform or recode variables are not well documented, and code used to conduct the actual analysis is

not provided. Most of the original authors try to be helpful, but occasionally they are entirely nonresponsive. My students often come away from this assignment frustrated, shocked, and rather disappointed by what they see as practices that undermine the credibility of the research they read. However, I also hope these experiences instill a resolve in them to make sure that their own work meets higher standards for data access and research transparency.

Finally, the principles of data access and research transparency should play a central role in our substantive courses. The typical graduate seminar engages and evaluates existing literatures relative to the theoretical and/or substantive topic at hand. The transparency and replicability of the research process used by authors should be a normal part of how young scholars are trained to evaluate existing studies. Similarly, student seminar papers should be evaluated, in part, on the transparency of the research methods used.

Adapting how we train students to incorporate data access and research transparency principles from the outset has many potential benefits. Researchers who learn to think about these issues at the start of their careers, and who see value in doing so at the start of each research project, will be better able to produce research consistent with these principles. Furthermore, meeting these goals should be easier for scholars trained this way from the outset—in fact, it will hopefully feel "automatic" or "natural" for students and scholars who experience and internalize this type of training. The best way to have an

represents, and even what each value for each bit of data represents. Thus, while a variable in a data set might consist of a column of zeros and ones, that data lacks meaning until you know that it was collected by a particular polling firm in October of 2012 via a telephone survey of registered voters, that the variable itself captures each respondent's intention to vote, and that a value of one indicates that the respondent intends to vote while a value of zero indicates that the respondent does not intend to vote. Such metadata is often described as a codebook for a data set, but modern data archiving links data and metadata directly rather than collecting metadata in a separate codebook.

Successful implementation of data access and research transparency principles requires careful attention to the production, documentation, and sharing of metadata. Metadata allows researchers to communicate information about their data sets as well as learn about other data sets. Metadata is the currency of data archives—it allows scholars to share, search, and discover what data exists and determine whether it might be of use to them. Sometimes researchers might need to limit access to the data itself, but they can still allow for the public distribution of metadata. Data might include identifying or other sensitive information that cannot be made public. Scholars might also want a period of time to exploit their data before sharing it with others. Proprietary restrictions might be needed on some data. However, scholars may be able to share metadata in each of these instances that would

The best way to have an enduring impact on how research is conducted in the future is to affect how researchers are trained in the present.

enduring impact on how research is conducted in the future is to affect how researchers are trained in the present.

There may be economies of scale that can be realized in this process. Common issues reach across social science disciplines, so departments might be able to collaborate. Numerous organizations with interdisciplinary training missions might also provide services. Some of the training necessary might be delivered online or through workshops rather than in traditional classes and seminars. In the end, however, the method of development and delivery is secondary to the more fundamental issue of deciding that data access and research transparency should be central elements of graduate education.

THE VALUE OF METADATA

Most discussions of data access and research transparency focus on the data itself. This focus has been amplified in recent years by both scholarly and public attention to the explosion of "Big Data." While raw data is essential, metadata is of equal importance. Metadata is best thought of as information about the data, or data about the data. Metadata provides meaning to data by describing it. Metadata includes information on who collected the data, when it was collected, where it was collected, how it was collected, and so forth.[8] Metadata also provides information on what each variable

provide for greater transparency for their research. Discovery of metadata by others might also create opportunities for collaboration or some other limited access to the data itself through a data use agreement. Thus, metadata is essential in its own right, but can also bridge the gap between complete data sharing and no transparency at all. Finally, training efforts regarding research transparency and data access should include explicit discussion of metadata from both a conceptual and applied perspective.

TURNING OBSTACLES INTO ADVANTAGES

NIH has required data sharing plans for grant proposals exceeding $500,000 since 2003. More recently, NSF released guidelines for meeting a new data management plan requirement. On May 9, 2013, President Obama issued an Executive Order, "Making Open and Machine Readable the New Default for Government Information," that requires the Office of Management and Budget to issue an Open Data Policy designed to make government data more widely available. These are just a few of the new policies pushing data openness and data sharing. Professional societies like the APSA are calling for more data sharing, and an increasing number of journals are adopting data sharing and replication policies. Such efforts certainly pose challenges in terms of privacy and data security,

and meeting these challenges will require researchers to work differently.

Individual scholars, departments, universities, disciplines and their professional associations, journals, publishers, and data archives can either resist these changes or they can lead them. They can wait to see what happens, or they can shape what happens. Change always creates disruptions, but it also presents innovators with opportunities. Those who incorporate data access and research transparency principles into their training programs and their own research practices, and those who invest time and effort into leading these efforts, are those who stand to gain.

Successful incorporation of data access and research transparency principles into the practice of research will ultimately make the entire research process, often called the research lifecycle, more efficient and productive. Scholars who are trained

seen an explosion in the use of open-source tools. The increased use of R for doing statistical analysis and of L^AT_EX for writing research reports alone has changed how many scholars and journals operate.[9] Open-source tools for analyzing data make it easier to provide open access to the data itself because it reduces the need for access to commercial software to exploit the data.

2. Increased use of online resources to distribute scholarly products. Numerous journals now publish accepted articles online before they are published in print. Publishers increasingly sell electronic access to journals to both individual and institutional subscribers, and more journals are providing online-only options for their subscribers. Given the economies of online publishing and the growing demand for online access to research materials, this trend is likely to continue. The good news is expanding online

> *Successful incorporation of data access and research transparency principles into the practice of research will ultimately make the entire research process, often called the research lifecycle, more efficient and productive.*

to track what they do from the beginning will make more efficient use of their research time. Those who seize the opportunity to develop training methods and new tools that advance these principles will attract more research dollars and more scholarly attention to their work. Journals that build easy-to-use systems for sharing all of the products of research should enhance the visibility of the work they publish. Departments and universities that reward scholars who exemplify data access and research transparency principles will be better able to attract and retain researchers committed to those principles.

Expanding efforts to promote greater data access and research transparency does add some additional burdens to researchers. However, effective implementation of these principles in training, workflow, and technical solutions will minimize those burdens and may increase opportunities for success by promoting visibility. While these pressures are emerging from a number of sources, a comprehensive evaluation of how social scientists conduct research and train future researchers offers a chance for healthy adaptation and reform. Again, the key to success is increasing the benefits associated with providing greater data access and research transparency while lowering the costs of doing so.

LOOKING FORWARD

A significant obstacle to successful adoption of data access and research transparency practices is the uneasy sense that changes will continue to happen faster than we can adapt. Thus, I close with some thoughts about four other trends affecting political science research that might be connected to greater data access and research transparency.

1. Increased use of open-source research tools. Research in the social sciences—particularly quantitative research—has

publication of research articles should make it easier to connect those studies to digital archives of the data such articles use.

3. Continued tension between data security and data openness. The demand for access to data of all types is increasing across all sectors of society, not just in the academic research community. Data about the attitudes, opinions, and behaviors of people—the bread and butter of social science research—is increasingly available, but it is posing new challenges for the protection of privacy and any other potential harm that might befall research subjects. The ability to mine data from online activities, use of digital devices, and so forth exposes research subjects to greater privacy risks. These broader ethical issues increase the pressure to make data access and research transparency more central elements of our graduate training programs.

4. Big Data. The term "Big Data" has gone from novel to overused very quickly. While definitions of Big Data differ, there is no denying the explosion of data about social processes that has become available. This trend seems destined to continue, meaning that any long-term solution to data access and research transparency concerns must consider how well it scales up to massively large data sets and the associated complex analytic methods used to analyze such data.

Fortunately, tools and ongoing research projects are focusing on addressing these trends. For example, tools like Sweave and knitr allow researchers to embed R code directly into a L^AT_EX document so that when the document is compiled in L^AT_EX, the R code is automatically executed.[10] This allows for research reports that include an analysis of the most recently available data to be generated on the fly. A byproduct of this

approach is that it embeds the replication code necessary to produce the reported analysis directly in the document, though researchers would need access to the uncompiled L^{A}T$_E$X document to see it.

Another tool currently available is a package that can be installed in R called Shiny.[11] Shiny allows researchers to create simple web applications that present output from R functions online. The code to produce those results can also be shown. Thus, research papers could be presented with interactive tables and/or figures that appear online and include the code used to produce them. This provides another mechanism whereby researchers can directly share more than just the final table or figure they wish to include in their paper—they can present the code that accesses the data necessary to produce that table or figure as well. If the underlying data is updated, the table or figure can be automatically updated as well.

The DVN includes several features designed to facilitate data sharing, data citation, and research replication. The DVN has extensive capabilities to help users produce quality metadata. It also includes some built-in analysis tools, a means of providing a unique digital identifier as part of a citation to data sets, and even the capacity to produce subsets of data and the corresponding code associated with any analysis a researcher might run within the DVN.

As director of the Odum Institute, and through involvement with APSA's DA-RT ad hoc committee, fortunately I have been engaged in some of the efforts directed at promoting greater data access and research transparency. Although I see many challenges, I strongly support the normative, ethical, and scientific values associated with greater research transparency. Still, the success of efforts designed to promote greater data access and research transparency will depend on whether they lower the costs and raise the benefits of adopting data access and research transparency principles, and whether we adapt our training programs so that these principles drive the establishment of updated norms about the proper conduct of research and dissemination of knowledge.

ACKNOWLEDGMENTS

NOTES

1. We can make this easier by encouraging scholars to frame their own research in terms of how it builds on existing studies rather than in terms of the problems with existing work.

2. The essay by John Ishiyama in this symposium devotes careful attention to this issue.

3. Those interested in learning more about the DVN should start here: http://thedata.org/.

4. The same applies for books and edited volumes as well.

5. See Vision (2010), the NERC Science Information Strategy Data Citation and Publication Project (http://ijdc.net/index.php/ijdc/article/view/208) and the DVN Integration project (http://projects. iq.harvard.edu/ojs-dvn) for examples. The Odum Institute also has a pilot project underway, supported in part by an ICPSR/Sloan Foundation Challenge Grant, to develop recommendations on integrating the article and data publication workflows.

6. As one who has served as a journal editor both with and without access to an online system, I can attest to this claim.

7. It is more likely that each would simply use links to the other.

8. Archives have developed a number of conventions and standards for the production of metadata. The DVN, for example, permits the generation of metadata that follows DDI, Dublin Core, FGDC and MARC standards.

9. The DVN is also open-source software.

10. Interested readers should consult the Sweave website: http://www.stat.uni-muenchen.de/~leisch/Sweave/ and the knitr website: http://yihui.name/knitr/.

11. Interested readers should start here: http://www.rstudio.com/shiny/.

REFERENCES

Gherghina, Sergiu, and Alexia Katsanidou. 2013. "Data Availability in Political Science Journals." *European Political Science* March: doi:10.1057/eps.2013.8.

Pienta, Amy M., George C. Alter, and Jared A. Lyle. 2010. "The Enduring Value of Social Science Research: The Use and Reuse of Primary Research Data." http://hdl.handle.net/2027.42/78307.

Piwowar, Heather A., Roger S. Day, and Douglas B. Fridsma. 2007. "Sharing Detailed Research Data is Associated with Increased Citation Rate." *PLoS ONE* 2 (3): doi:10.1371/journal.pone.0000308.

Vision, Todd J. 2010. "Open Data and The Social Contract of Scientific Publishing." *Bio-Science* 60 (5): 330–331.

Replication, Research Transparency, and Journal Publications: Individualism, Community Models, and the Future of Replication Studies

John Ishiyama, *University of North Texas*

Recently, the importance of research transparency via replication studies has been greatly discussed in most of the social sciences, political science included. Indeed, as Gherghina and Katsanidou (2013) and Freese (2007) note, to some extent, the discussion has been prompted by the tremendous changes in publishing in the past decade or so. With the enormous expansion in data availability and instant publication made possible by the Internet, there now are many opportunities to verify the findings presented in the discipline's major journals. "Replication, replication" has not only become the mantra for political science, but for economics, psychology, and quantitative sociology as well. These developments opened a debate on how to best "guard the high standards or research practice and allow for the maximum use of current knowledge for the further development of science" (Gherghina and Katsanidou 2013, 1; for similar sentiments see King 1995).

Some scholars who advocate greater research transparency via replication studies have provided guidelines for what should be included in publicly available replication files. For instance, Gary King (2003) has proposed a checklist of what should be included when making data available for replication. These items include the original data, the specialized software that was used, syntax files, extracts of existing data files, and comprehensive documentation to explain how to reproduce the exact output presented in the published work. Further, several journals in political science and international relations have followed these guidelines and sought to make data available for replication studies.

Certainly there are many advocates of promoting replication in political science, however, in this essay I focus on two questions that the move toward research transparency and replication raise for journal publication in political science. First, what are the implications for the journals with the shift from an "individual model" of responsibility for the provision of replication data to a more "social policy" or "community model'? Second, and perhaps more important, where should studies that replicate existing works be published?

THE DEBATE OVER INDIVIDUAL AND COMMUNITY RESPONSIBILITY

Regarding the first question of responsibility, generally speaking, there is the distinction made between the "individualistic" model and the "social" or "community" model of collective responsibility (Freese 2007; King 2006). On the one hand, the individualistic model holds that the primary responsibility for making data available for replication purposes lies with the individual author.[1] On the other hand, the social or community policy makes the provision of replication data part of the publication process (and required by journals as requirement for publication). In this case the journal, as a representative of the scholarly community, is the responsible to make sure that data for replication purposes is provided to that community.

Of the two approaches, the literature on replication clearly favors the adoption of the community model (or social policy) over the individualistic model. Indeed, the community model has some important advantages. As Freese (2007) notes, if a policy is enforced by the journals, readers can fully expect that the data are already provided for replication. In contrast, in the individualistic model, the reader would have to trust that the author will provide the reader with data for replication on request. Further, the community model of replication data provision guarantees that such data would be preserved over time in a reproducible format. The individualistic model relies on the individual scholar's ability to preserve such data, which may or may not happen. In other words the "social policy seeks to decouple the content of articles from the contingencies of authors' futures" (Freese 2007, 156)

Further, as Freese (2007, 156) argues, a certain egalitarianism is promoted by the community model as it minimizes the "degree to which status and social networks affect access to materials necessary to verify, learn from, and build off of others' work" (see also King 2006). In the individualistic model, in contrast, there is the possibility for the selective release of data. In other words, replication data would be more readily provided to notable faculty or from certain elite institutions, than to junior faculty, graduate students, or faculty members from less prestigious institutions.

Although the adoption of a community model for replication access generally benefits the scholars in political science, the adoption of such policies that increase access to replicable data also directly benefits the journals. Certainly, as Freese (2007, 156) notes, such an approach "also increases the extent to which articles that command scarce journal space are instructive to other researchers by allowing interested others to see more details of how exemplary work was done." However, in

addition, there is something also very practical regarding promoting widespread proliferation of replication studies: the prospect that work will be replicated promotes greater scholarly honesty in research. The pressure to produce "positive results" provides all sorts of incentives for "cooking" or "massaging" the results, and, in the worst case, for falsification of findings. Knowing that their works will be replicated (and perhaps even more importantly that this replication will be available for public scrutiny) holds the original authors accountable for their work (which is also true, I would imagine, for qualitative work as well), thus acting as a deterrent to such irresponsible behavior. The adoption of such a standard certainly will not solve all issues regarding academic honesty or prevent the search for "positive results" (certainly the registration of research designs prior to the conduct of a project would also dis-incentivize such behaviors) but it would be a big step in the right direction.

Note, however, that the move toward the adoption of replication standards and data transparency has not been without its critics (Gherghina and Katsanidou 2013). For instance, James Gibson (1995) has argued strongly against the introduction of journal-enforced replication standards as implied by favoring the community model, suggesting that such a move would lead to a focus on minor methodological "trivia" as

of all replication studies, but this is beyond the capacity (and currently the willingness) of most journals.

Although the literature has focused largely on advocating the provision of replication data, and the obvious benefits for the scholarly community, much less analysis has been done empirically on the current state of the discipline and how political science compares with other fields in the social sciences.

WHAT IS THE CURRENT STATE OF THE FIELD AMONG JOURNALS IN THE SOCIAL SCIENCES?

Little empirical work has examined the state of the discipline regarding how journals deal with the issue of the provision of replication data. However, a recent very important exception is an article in *European Political Science* by Sergiu Gherghina and Alexia Katsanidou (2013). Surveying journal websites, and then following up with a survey of editors, in their study of 120 political science and international relations journals, the authors found that only 19 journals had any policies regarding replication (but importantly, most all of the high-impact journals and general journals had such policies in place).

Although this may seem a remarkably low proportion of journals, the lack of emphasis on provision of replication data is not limited to political science. Quantitative sociologists also have long lamented the lack of the availability of replication

A potential concern rising from an emphasis on replication studies, is that such an emphasis, if endorsed by the discipline's major journals, will incentivize "witch hunts" and an effort to "slay giants" as a career pursuit. Perhaps this can be allayed by careful review of all replication studies, but this is beyond the capacity (and currently the willingness) of most journals.

opposed to theory and a minimization of the value of the analysis of large secondary data sets in favor of small original ones (Gibson 1995, 475).

Another concern that potentially arises from the journal-enforced replication standards is that it may lead to poorly conducted replication studies that are submitted to the journals. As Funder (2013) contends in an editorial "Does 'Failure to Replicate' Mean Failed Science?", although some egregious cases in psychology may cause alarm, such fraud is actually very rare and "focusing on them too tightly can be misleading." There are many reasons why replications fail—the replication study may not follow the exact methods used by the original research project; or the replicator lacks the necessary skills to replicate the original study; or the original finding simply may have been a "lucky accident." Further, scholars work very hard to work through the "chaos" of social and political reality and often are quite eager to make their results public. Although at times they may be too eager to report results, scholars' reputations and careers are on the line. A potential concern rising from an emphasis on replication studies, is that such an emphasis, if endorsed by the discipline's major journals, will incentivize "witch hunts" and an effort to "slay giants" as a career pursuit. Perhaps this can be allayed by careful review

data in the leading sociology journals (see Freese 2007, for a valuable overview of the situation in sociology). In economics, largely as the result of a series of studies that reported dismal rates of both author cooperation and lack of reproducible results (Dewald, Thursby, and Anderson 1986; McCullough, McGeary, and Harrison 2006; McCullough and Vinod 2003), the official journals of the American Economic Association that publish original empirical research now have an extensive policy regarding the availability of data and materials for replication.

However, perhaps the greatest effort to address the issue of replication has occurred in psychology, and in many ways psychology is taking the lead in promoting data access and replication studies in the social sciences (Funder 2013). Although a long tradition of experimental replication exists in the field, nonetheless, there have been remarkably low levels of cooperation in data sharing. A study of 141 articles in American Psychological Association journals—whose stated policy is similar to many political science journals and sociology journals in putting the responsibility of data availability on the authors found only 27% compliance with repeated requests for data for verification purposes (Wicherts et al. 2006). More recently, concern over the falsification of results is growing as well as a call for the provision of data for

<table>
<thead>
<tr><th colspan="3">*Table 1*
Three Models of Journal Replication Data Management</th></tr>
<tr><th>MODEL</th><th>WHO IS RESPONSIBLE FOR DATA PROVISION?</th><th>WHO IS RESPONSIBLE FOR PROVISION ENFORCEMENT?</th></tr>
</thead>
<tbody>
<tr><td>*Journal Responsibility Model*</td><td>Scholar provides journal with data, journal is responsible for request for replication data.</td><td>Journal provides data on request.</td></tr>
<tr><td>*Journal Certification Model*</td><td>Individual scholar is responsible for provision of replication data on request.</td><td>Journal certifies that individual has provided accessible data before publication.</td></tr>
<tr><td>*Trust Model*</td><td>Individual scholar is responsible for provision of replication data on request.</td><td>Journal trusts individual author(s) to provide data.</td></tr>
</tbody>
</table>

reproducibility purposes (the recent case of discredited Dutch social psychologist Diederik Stapel has highlighted these concerns).[2] So great has this concern become in psychology, that a group of psychologists have launched "the Reproducibility Project" as apart of the "Open Science Framework," which aims to replicate the results from leading psychological journals that appeared in articles in 2008 (*Psychological Science*, the *Journal of Personality and Social Psychology*, and the *Journal of Experimental Psychology: Learning, Memory, and Cognition*) (see http://chronicle.com/blogs/percolator/is-psychology-about-to-come-undone/29045).

Thus, the problem for providing reproducible data for replication is not a challenge only facing political science, but most of the social sciences.

INDIVIDUAL VERSUS COMMUNITY RESPONSIBILITY IN POLITICAL SCIENCE JOURNALS

What of the issue of individual versus community responsibility for the provision of replication data? A closer look at the journals in political science and international relations reveals a mixed picture in terms of who is responsible for providing access to data for replication. As Gherghina and Katsanidou (2013) point out, most of the journals in political science and international relations do not have a policy regarding replication. Of the 19 that do, most emphasize individual responsibility for the provision of data for replication purposes. Some very important exceptions exist, particularly the recent changes adopted by the *American Journal of Political Science*, and the policies several leading international relations journals as well.

Thus far I have only discussed the distinction made in the current literature between individual and community-based models of responsibility for the providing data for replication in terms of *either* the journals provide access to data or the individual authors do. Perhaps it would be more useful to frame the choices in terms of *provision of data* (or who is responsible for holding the replication files and making them available on request) and *enforcement of provision* (or who makes sure that the data are actually accessible).

Table 1 illustrates three basic models of replication files management, based on these two dimensions. First is what I label the *Journal Responsibility Model* (JRM, which is a form of community provision), where the journal requires that data is provided to the journal prior to publication of an article (and can be stored either by the journal or at a community site such

as *dataverse*) which then makes it available on request to scholars who seek to replicate the findings of the study. The journal naturally enforces provision of the data. A second model, the *Journal Certification Model* (JCM) is also a form of community provision, but is different from the *Journal Responsibility Model* in that the individual author(s) are responsible for holding the data and making it available (perhaps on the scholar's website) but the journal enforces provision by requiring some form of certification that data is accessible prior to publication of the article (a variation of this model might be that the journal "requires" public provision, but does not enforce this requirement). In the third model, the *Trust Model*, the author(s) are responsible for provision of the data and the journal trusts that the author(s) will provide the data on request.

Generally, the norm by the journals has been to emphasize the individual's responsibility of providing replication data when requested, and that the journals will generally trust that this is done (or the *trust model*). For instance, the *American Political Science Review* emphasizes this when the instructions to the authors asks authors that if

> your manuscript contains quantitative evidence and analysis, you should describe your procedures in sufficient detail to permit reviewers to understand and evaluate what has been done and—in the event the article is accepted for publication—to permit other scholars to replicate your results and to carry out similar analyses on other data sets.... In addition, authors of quantitative or experimental articles are expected to address the issue of data availability. You must normally indicate both where (online) you will deposit the information that is necessary to reproduce the numerical results and when that information will be posted (such as "on publication" or "by [definite date]"). You should be prepared, when posting, to provide not only the data used in the analysis but also the syntax files, specialized software, and any other information necessary to reproduce the numerical results in the manuscript.

Similar language regarding the provision of data for replication purposes is available in the *Journal of Politics*

> Authors of quantitative papers published in the JOP must address the issue of data availability in *Footnote 1* of their paper. Authors are expected to indicate both where (online) they will deposit the information necessary to reproduce their numerical results and when that information will be posted. Authors should include

not only the data used in the analysis but also the syntax files, specialized software, and any other information necessary to reproduce the numerical results in the manuscript. A statement explaining why the data or other critical materials used in the manuscript cannot be shared, or justifying their embargo for a limited period beyond publication may fulfill this requirement. However, we strongly encourage our authors to comply with the spirit of this policy and embrace the scientific norms of professional accountability and openness.

Although the guidelines include the checklist offered by King (2003), in both cases it is clearly the author's responsibility to provide data, not the journal's. There are no specific measures to ensure that the data is actually provided, other than that an expectation is expressed that authors do so. Neither journal currently provides a site for the making replication files available for its published pieces.

In contrast, the *American Journal of Political Science* has recently moved in the direction of community provision of replication files, in terms of both submission of data and certification that such data will be accessible if not submitted prior to publication of the article. The journal requires that on acceptance for publication the "manuscript will not be published unless the first footnote explicitly states where the data used in the study can be obtained for purposes of replication and any sources that funded the research." Further, and perhaps most important, the journal provides a site for storage of all replication files at the "AJPS Data Archive on Dataverse."

Several major international relations journals, particularly those associated with the International Studies Association (ISA) have generally followed the *Journal Responsibility Model* and the *Journal Certification Model* in that they require the provision of replication files as a condition for publication, and these files are posted publicly by the journals. This was a direct result of a symposium on "Replication in International Studies Research" organized by one of the association's journals, *International Studies Perspectives*, in 2003. The symposium was derived from a set of papers that had been presented at the 2002 International Studies Association Meeting in New Orleans. As a result of these efforts, four leading international relations journals adopted a single common replication policy (James 2003; Gleditsch et al. 2003a; Gleditsch et al. 2003b)—these included *International Studies Quarterly, Journal of Peace Research, Journal of Conflict Resolution*, and *International Interactions*.

One of these journals was the flagship journal of the ISA, the *International Studies Quarterly (ISQ)*, whose submission guidelines clearly state the requirement that authors make "their data . . . . fully accessible. If the data in question are not already publicly archived, authors will be required to certify that the data are readily available to others. Requests for copies of the data must be addressed to the author or authors, and not the offices of *ISQ*." Thus, there is no requirement that data be deposited with *ISQ* as long as the author can document that it is publicly archived elsewhere. If not, data is archived with the journal and made public on the ISA's website at http://www.isanet.org/Publications/ISQ/ReplicationData.aspx.

As these examples illustrate, there is considerable variation in the implementation of replication policies by journals in political science (and to a lesser extent international relations). Many journals do not have any policies to speak of (as clearly indicated by the work of Gherghina and Katsanidou 2013). Even those journals that do, only a few have embraced the community-based model of requiring submission of replication files *prior* to publication of an article in a journal. Why have journals been slow to adopt a community-based standard?

One possible reason for the hesitancy is the lack of space to store replication files. This is probably more true for specialized journals that do not have the resources of the major general journals that are supported by subsidies from major academic presses. However, insufficient storage space may become less of a problem with the availability of such storage sites as "Dataverse" or by storage sites made available by professional associations (such as the ISA).

A more vexing problem is what to do with nonquantitative pieces that appear in the journals. Indeed, in many journals, including the major ones, the emphasis on qualitative and/or normative work is increasing, which does not lend itself as easily to storage and access (and do not necessarily follow the protocol provided by King 2003). The major journals are imprecise about data provision and enforcement and mostly leave the provision of qualitative data entirely up to the authors. Thus the APSR states:

> … authors of qualitative, observational, or textual articles, or of articles that combine such methods with quantitative analysis, should indicate their sources fully and clearly enough to permit ready verification by other scholars—including precise page references to any published material cited and clear specification (e.g., file number) of any archival sources. Wherever possible, use of interactive citations is encouraged. Where field or observational research is involved, anonymity of participants will always be respected; but the texts of interviews, group discussions, observers' notes, etc., should be made available on the same basis (and subject to the same exceptions) as with quantitative data. (see http://www.apsanet.org/content_43805.cfm)[3]

However, as indicated in several pieces in this *PS* symposium (particularly the contributions by Elman and Kapizsewski, and Moravcsik) new standards and new ideas for the provision of qualitative data for research transparency purposes are being developed. Thus, the major journals soon should be in a position to enact some of these recommended standards.

PUBLICATION VENUES FOR REPLICATION STUDIES?

Perhaps a more important issue, at least from the perspectives of the journals (which has not received nearly as much attention in the literature) is where replication studies should be published. If the prospects of public replication of published work is to deter scholarly dishonesty or misrepresentation of results, identifying a venue for the publication of such work should be a central part of any discussion of the adoption of replication policies in political science. Simply providing access to data is not enough—an outlet for the publication of such material provides an incentive for scholars to engage in such an often time-consuming activity with little obvious rewards.

The editors of the *APSR* have been discussing this issue for some time. In many ways this was prompted by several recent exchanges we had with a scholar who had obtained the replication data from the authors of a manuscript that had appeared in an earlier issue of the *Review* (in 2010, prior to the University of North Texas' team taking the reins of the journal). After obtaining the replication data from the authors of the original piece (with the editors' help) they proceeded to attempt to replicate the results, but were unable to do so. The authors notified us and asked where to publish such a replication study. Our policy at the *APSR* (which was also the policy of all of our predecessors and the policy of most major journals in the social sciences as well) is not to publish works that are only replication studies because they do not represent the kind of *original* work we publish in the *Review*.

There are very good reasons for *APSR*'s policy, and we strongly believe in continuing it. We do believe, however, that a very good point was made. A venue for the publication of replication studies is necessary, especially the discipline aspires to raise the degree of scientific rigor in the field. However, as editors of the *APSR* we are also reluctant to publish such studies in the *Review*, because this would open up a "cheap" way for authors to have their work published in the *APSR*, and every Tom, Dick, and Harriet (pardon the expression) could potentially seek to replicate some study, just to get published in the *Review*. Most all other major journals in the field, we believe, do not to publish solely replication studies (certainly this is true of *APSR*, *AJPS* and *JOP*, as well as the major international relations journals).

Certainly in the past occasional "Forums" have been published in the *Review*, and in other journals, as well. This potentially allows for the incorporation of such replication studies in a rebuttal and a rejoinder, however, these instances are too rare to address the general issue. No current venue provides for the publication of replication studies of pieces that appear in *APSA* journals, that appear in an APSA venue (some replications of *APSR* articles appear in journals outside of APSA, but not in an APSA publication). If we are serious about promoting research transparency and scholarly integrity via access to replication files, we must also, as a community, provide a venue for this material to be made public (and published). Given the challenges associated with publishing replication attempts, researchers now have little incentive to conduct such studies.

What are some ways to provide such publication venues? One model is offered by psychology. The Association for Psychological Science (APS) has provided a special section in one of the society's journals dedicated to the production of replication reports. Note that replication studies rarely appear in psychology journals. The new *Registered Replication Reports* article type in *Perspectives on Psychological Science* seeks to provide an outlet for work that replicates research in psychology. The journal argues that:

- psychological science should emphasize findings that are robust, replicable, and generalizable;
- direct replications are necessary to estimate the true size of an effect;

- well-designed replication studies should be published regardless of the size of the effect or statistical significance of the result; and
- traditional psychology journals do not have the space or inclination to publish such reports (see http://www.psychologicalscience.org/index.php/news/releases/initiative-on-research-replication.html).

Note that *Perspectives on Psychological Science* (although a highly ranked journal) is not a general research journal. Rather, its purpose and function is similar to *PS* and *Perspectives on Politics* in political science, and *International Studies Perspectives* and the *International Studies Review* in international relations. As such, it publishes "reports and articles, including broad integrative reviews, overviews of research programs, meta-analyses, theoretical statements, book reviews, and articles on topics such as the philosophy of science, opinion pieces about major issues in the field, autobiographical reflections of senior members of the field, and even occasional humorous essays and sketches." To follow this model would require a special section of *PS* reserved for replication studies

Another second minimalist alternative would be to provide an electronic "blog like" venue for the publication of replication studies, something like the "Monkey Cage" a very popular blog/newsletter that is read by thousands of political scientists (and policy makers) throughout the world. Certainly this would make replication findings more public, and require considerably less space in an existing journal (and less resources than a new journal), and certainly could be seen as a deterrent on scholarly dishonesty. However would a blog carry the same prestige come tenure and promotion time as published in a peer-reviewed publication? This strategy would not provide as strong an incentive scholars to conduct replication studies, and without such studies the deterrent effect of replication would be minimized.

Another third model would be to offer "publication" of replication studies by the major journals in the discipline, but to print those replication studies in an online supplement directly linked to the articles that appear in the journals. Journals could highlight those articles that have been replicated multiple times, providing an important service to readers, and a greater reward for better work.

A fourth model is to create an entirely new publication. This includes considering a new APSA publication (or perhaps part of a proposed new publication). Currently the association, largely as the result of the efforts of the APSA immediate past president Jane Mansbridge, has begun to assess the current array of journals and to plan for any additional journals for publication as is necessary for the discipline. Such a journal, if launched by the association, could have, as one of its core missions, the publication of replication reports, in addition to other functions.

In short, the APSA should consider potential alternative venues for the publication of replication studies (or perhaps "forums" or debates) of pieces that appear in APSA journals. Now it is not exactly clear how this should be done, if it could be done online, if it requires an editorial team, what the relationship would be with the existing APSA journals, and how would this be related to Cambridge University Press, but if we

are to move forward as a discipline, we *must* have some venue available for the publication (or at least making public) such studies.

CONCLUDING REMARKS

This article argues that the move toward the adoption of replication policies by the major journals in political science raises two issues. One, who should be responsible for the provision of replication materials to the scholarly community? And, two, where should these replication studies be published?

First, the move toward a community or social policy model is preferable to the individualistic policies adopted by most journals (either in the form of the *Journal Responsibility Model* or the enforced *Journal Certification Model*) but this raises issues of space and storage (particularly regarding the JRM) as well as what to do with qualitative and normative work and other forms of research that are published in many general political science journals. Only providing replication materials for quantitative studies would not only be incomplete, but would send the signal that only quantitative studies should be externally validated, and that other, less important work need not be. Clearly, this is not the message that the major journals should communicate to the scholarly community. The other contributions in this symposium highlight how the journals might more effectively begin to deal with issues of data access and research transparency for qualitative work.

Second, an outlet for the publication of replication studies that appear in APSA journals is needed (although not necessarily exclusively on articles that appear in APSA journals), that the APSA should publish. This might involve one of the four alternative approaches discussed earlier, or perhaps another approach. Whatever the case, this is something that should be part of the discussion of replication and research transparency that has not, in my view, been adequately addressed. ■

NOTES

1. This is not to suggest that individuals who are responsible for provide replication data are not responding to group norms emanating from a scholarly community. It means that the primary *responsibility* for providing data lies with the author, not the journal.

2. The Stapel case is not the only recent controversy in psychology that has increased the call for more replication studies. For a discussion of other cases see Roediger (2012).

3. As for the *AJPS*, the guidelines do not speak directly to the issue of qualitative data at all, although the guidelines speak of "supporting information" and such material must be "made ready for permanent posting" but manuscripts without data or SI are exempt." http://www.ajps.org/manu_guides.html

REFERENCES

Dewald, William G., Jerry G. Thursby, and Richard G. Anderson. 1986. "Replication in Empirical Economics: The Journal of Money, Credit, and Banking Project." *The Journal of Money, Credit, and Banking* 76: 587–603.

Freese, Jeremy. 2007. "Replication Standards for Quantitative Social Science: Why Not Sociology?" *Sociological Methods and Research* 36: 153–72.

Funder, David. 2013. "Does 'Failure to Replicate' Mean Failed Science?" *Live Science* at http://www.livescience.com/32041-revisiting-science-studies.html (accessed September 30, 2013).

Gherghina, Sergiu, and Alexia Katsanidou. 2013. "Data Availability in Political Science Journals." *European Political Science*. Advance online publication, March 1; doi:10.1057/eps.2013.8.

Gibson, James L. 1995. "Cautious Reflections on a Data Archiving Policy for Political Science." *PS: Political Science and Politics* 28: 473–76.

Gleditsch, Nils Petter, Patrick James, James L. Ray, and Bruce Russett. 2003a. "Editors' Joint Statement: Minimum Replication Standards for International Relations Journals." *International Studies Perspectives* 4: 105.

Gleditsch, Nils Petter, C. Metelits, and H. Strand. 2003b. "Posting Your Data: Will You Be Scooped or Will You Be Famous?" *International Studies Perspectives* 4: 89–97.

James, Patrick. 2003. "Replication Policies and Practices in International Studies Quarterly." *International Studies Perspectives* 4: 85–88.

King, Gary. 1995. "Replication, Replication." *PS: Political Science and Politics* 28: 443–99.

King, Gary. 2003. "The Future of the Replication Movement." *International Studies Perspectives* 4: 100–05.

King, Gary. 2006. "Publication, Publication." *PS: Political Science and Politics* 39: 119–25.

McCullough, B. D., Kerry Anne McGeary, and Teresa D. Harrison. 2006. "Lessons from the JMCB archive." *Journal of Money, Credit, and Banking* 38: 1093–107.

McCullough, B. D., and H. D. Vinod. 2003. "Verifying the Solution from a Nonlinear Solver." *American Economic Review* 93: 873–92.

Roediger, Henry. 2012. "Psychology's Woes and a Partial Cure: The Value of Replication." *APS Observer* 25 (2) at http://www.psychologicalscience.org/index.php/publications/observer/2012/february-12/psychologys-woes-and-a-partial-cure-the-value-of-replication.html (accessed September 30, 2013).

Wicherts, Jelte M., Denny Borsboom, Judith Kats, and Dylan Molenaar. 2006. "The Poor Availability of Psychological Research Data for Reanalysis." *American Psychologist* 61: 726–28.