Qualitative & Multi-Method Research

Contents

Letter from the Editors

It is an honor to take up the reins as editors of *Qualitative and Multi-Method Research*. For 12 years, this newsletter has offered incisive discussion and spirited debate of the logic and practice of qualitative and multi-method research. It has been a venue for critical reflection on major new publications and for the presentation of novel arguments, long before they hit the journals and the bookshelves. Among the newsletter's most important hallmarks has been the inclusion of—and, often, dialogue across—the wide range of scholarly perspectives encompassed by our intellectually diverse section. We are grateful to our predecessors John Gerring, Gary Goertz, and Robert Adcock for building this newsletter into a distinguished outlet, and delighted to have the opportunity to extend its tradition of broad and vigorous scholarly engagement.

In this issue, we present a symposium on an issue that has been a central topic of ongoing conversation within the profession: research transparency. The symposium seeks to advance this conversation by unpacking the meaning of transparency for a wide range of qualitative and multi-method research traditions, uncovering both considerable common ground and important tensions across scholarly perspectives.

We also present here the first installment of our new Journal Scan, based on a systematic search of nearly fifty journals for recent methodological contributions to the literature on qualitative and multi-method research. Building on the "Article Notes" that have occasionally appeared in previous issues, the Journal Scan is intended to become a regular element of *QMMR*. We look forward to hearing your thoughts about this newsletter issue and your ideas for future symposia.

Tim Büthe and Alan M. Jacobs

Transparency in Qualitative and Multi-Method Research

Introduction to the Symposium

Tim Büthe
Duke University

Alan M. Jacobs
University of British Columbia

Research transparency has become a prominent issue across the social as well as the natural sciences. To us transparency means, fundamentally, providing a clear and reliable account of the sources and content of the ideas and information on which a scholar has drawn in conducting her research, as well as a clear and explicit account of how she has gone about the analysis to arrive at the inferences and conclusions presented—and supplying this account as part of (or directly linked to) any scholarly research publication.

Transparency so understood is central to the integrity and interpretability of academic research and something of a "meta-standard." As Elman and Kapiszewski suggest, transparency in social science is much like "fair play" in sports—a general principle, the specific instantiation of which will depend on the particular activity in question.[1] As we will discuss in greater detail in our concluding essay, research transparency in the broadest sense also begins well before gathering and analyzing empirical information, the research tasks that have been the focus of most discussions of openness in Political Science.

As a consequence of a number of developments[2]—

including very importantly the work of the American Political Science Association's Ad Hoc Committee on Data Access and Research Transparency (DA-RT) and the resulting incorporation of transparency commitments into the 2012 revision of the APSA *Ethics Guide*[3]—the issue of research transparency has recently moved to the center of discussion and debate within our profession. We share the assessment of many advocates of research transparency that the principle has broad importance as a basis of empirical political analysis, and that most of the concerns that have motivated the recent push for greater openness are as applicable to qualitative and multi-method (QMM) research as to quantitative work. Yet, however straightforward it *might* be to operationalize transparency for research using statistical software to test deductively derived hypotheses against pre-existing datasets, it is not obvious—nor simple to determine—what transparency can and should concretely mean for QMM research traditions. The appropriate meaning of transparency, moreover, might differ across those traditions, as QMM scholars use empirical information derived from a great variety of sources, carry out research in differing contexts, and employ a number of different analytical methods, based on diverse ontological and epistemological premises.

To explore the role of research transparency in different forms of qualitative and multi-method research, this symposium brings together an intellectually diverse group of scholars. Each essay examines what transparency means for, and demands of, a particular type of social analysis. While some essays focus on how principles of research transparency (and data access) might be best and most comprehensively achieved in the particular research tradition in which the authors work, others interrogate and critique the appropriateness of certain transparency principles as discipline-wide norms. And some advocate forms of openness that have not yet featured prominently on the transparency agenda. We hope that the discussions that unfold in these pages will contribute to a larger conversation in the profession about how to operationalize a shared intellectual value—openness—amidst the tremendous diversity in research practices and principles at work in the discipline.

[1] Elman and Kapiszewski 2014, 46n3.

[2] To better understand the broader context of the calls for greater transparency in Political Science, we have examined recent debates over research transparency and integrity in disciplines ranging from Life Sciences (medicine and associated natural sciences, as "hard" sciences that nonetheless gather and analyze empirical data in a number of ways, including interviews, the analysis of texts, and case studies, while facing ethical and legal requirements for confidentiality and generally the protection of human subjects) to History (as a discipline on the borderline between the social sciences and the humanities, with a long tradition of dealing with uncertainty and potential bias in sources, though usually without having to deal with concerns about human subjects protection). While a detailed discussion of developments in those other disciplines is beyond the scope of

this symposium, we have found the consideration of developments in other fields helpful for putting developments in Political Science in perspective. We were struck, in particular, by how common calls for greater research transparency have been across many disciplines in recent years—and by similarities in many (though not all) of the concerns motivating these demands.

[3] APSA 2012, 9f. We refer to changes to the "Principles for Individual Researchers" section (under III.A.). Specifically, sections III.A.5 and III.A.6 were changed. Our understanding of the DA-RT Initiative and the 2012 revision of the *Ethics Guide* has benefited from a wealth of background information, kindly provided by current and past members of the the APSA staff, APSA Council, and APSA's Committee on Professional Ethics, Rights, and Freedoms, and on this point especially by a phone interview with Richard Johnston, ethics committee chair from 2009 to 2012, on 20 July 2015. These interviews were

In the remainder of this introductory essay, we first contextualize the symposium by presenting the rationales underlying, and key principles to emerge from, the recent push for transparency in Political Science. We then set out the questions motivating the symposium and briefly introduce the contributions that follow.

The Call for Greater Transparency: Rationales and Principles

The most prominent push for greater transparency in Political Science has emerged from the DA-RT initiative.[4] Scholars associated with this initiative have provided the most detailed conceptualization and justification of research transparency in the discipline. The DA-RT framework has also served as an important reference point for discussions and debates in the discipline and in this symposium, including for authors who take issue with elements of that framework. We therefore summarize here the central components of the DA-RT case for research transparency.

DA-RT's case for greater transparency in the discipline has been grounded in an understanding of science as a process premised on openness. As Lupia and Elman write in their introductory essay to a January 2014 symposium in *PS: Political Science and Politics*:

> 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 scholars to rigorously evaluate.[5]

Research openness, understood in this way, promises numerous benefits. Transparency about how findings have been generated allows those claims to be properly interpreted. Clarity about procedures, further, facilitates the transmission of insights across research communities unfamiliar with one another's methods. For scholars, transparency is essential to the critical assessment of claims, while it also makes social scientific research more useful to practitioners and policy-makers, who will often want to understand the foundations of findings that may inform their decisions.[6]

Transparency can be understood as a commitment to addressing, comprehensively, three kinds of evaluative and in-

terpretive questions, corresponding to the three principal forms of transparency as identified in the APSA *Ethics Guide*: production transparency, analytic transparency, and data access.[7]

First: *How has the evidence been gathered?* There are always multiple ways of empirically engaging with a given research topic: different ways of immersing oneself in a research context, of searching for and selecting cases and evidentiary sources, of constructing measures. Moreover, different processes of empirical engagement will frequently yield different collections of observations. Alternative ways of engaging with or gathering information about a particular case or context could yield differing observations of, or experiences with, that case. And, for research that seeks to generalize beyond the instances examined, different ways of choosing cases and research contexts may yield different *kinds* of cases/contexts and, in turn, different population-level inferences. DA-RT advocates argue that understanding and evaluating scholarly claims thus requires a consideration of how the particular interactions, observations, and measures upon which a piece of research rests might have been shaped by the way in which empirical information was sought or gathered.[8] Readers can only make such an assessment, moreover, in the presence of what the revised APSA ethics guidelines call "production transparency," defined as "a full account of the procedures used to collect or generate the data."[9] One might more broadly conceive of production transparency as a detailed account of the process of empirical engagement—whether this engagement is more immersive (in the mode of ethnography) or extractive (in the mode, e.g., of survey research or document-collection).

Second: *How do conclusions or interpretations follow from the empirical information considered?* What are the analytic or interpretive steps that lead from data or empirical engagement to findings and understandings? And what are the analytic assumptions or choices on which a scholar's conclusions depend? Greater "analytic transparency" is intended to allow research audiences to answer these questions, central to both the assessment and interpretation of evidence-based knowledge claims.[10] In the quantitative tradition, analytic transparency is typically equated with the provision of a file containing a list of the commands issued in a statistical software package to arrive at the reported statistical findings, which in principle allows for a precise specification and replication of the analytic steps that researchers claim to have undertaken. In the qualitative tradition, analytic transparency typically means making verbally explicit the logical steps or interpretive processes linking observations to conclusions or understandings. In process tracing, for instance, this might include an explicit discussion of the compatibility or incompatibility of individual pieces of evidence with the alternative hypotheses

---

conducted in accordance with, and are covered by Duke University IRB exemption, protocol # D0117 of July 2015.

[4] The initiative was launched in 2010 under the leadership of Arthur Lupia and Colin Elman and has *inter alia* resulted in the 2012 revision of the APSA *Ethics Guide* and the October 2014 "DA-RT Statement," which by now has gathered the support of the editors of 26 journals, who have committed to imposing upon authors three specific requirements: to provide for data access, to specify the "analytic procedures upon which their published claims rely," and to provide references to/for all pre-existing datasets used (see DA-RT 2014).

[5] Lupia and Elman 2014, 20.

[6] Lupia and Elman 2014, 22.

[7] Our discussion here draws partly on Lupia and Elman 2014 and Elman and Kapiszewski 2014.

[8] Lupia and Alter 2014, 57.

[9] APSA 2012, 10.

[10] See, e.g., APSA 2012, 10; Lupia and Alter 2014, 57; Lupia and Elman 2014, 21f.

under examination.[11]

Third: *What is the relationship between the empirical information presented in the research output and the broad evidentiary record?* One question here concerns the faithful representation of the evidence gathered by the researcher: Does or did the source (i.e., the document, informant, interviewee, etc.) say what the researcher claims that it says? A second, more difficult question is whether the empirical information employed in the analysis or presented in the writeup is representative of the full set of relevant empirical information that the researcher gathered, consulted, or otherwise had readily available.[12] "Data access" is intended to address these questions by giving readers direct access to a broad evidentiary record. Data access also allows other researchers to replicate and evaluate analytic procedures using the original evidence. Data access might include, for instance, making available full copies of, or extended excerpts from, the primary documents examined, the transcripts of interviews conducted, or the quantitative dataset that was analyzed.[13]

Transparency Deficits and their Importance to QMM Research

The most prominent concern motivating the push for transparency in many of the natural sciences has been outright research misconduct—specifically the fabrication of data or other source information or the falsification of findings.[14] A 2011 study by Fang and Casadevall, for instance, showed that the annual number of retractions of articles indexed in PubMed has over the past 25 years increased at a much faster rate than the number of new PubMed entries;[15] and the share of those retractions attributed to research misconduct has been large and growing.[16] In History, the more prominent problem

appears to be plagiarism,[17] though this field has been rocked by its own fabrication scandals.[18] Such developments have contributed to the push for greater research transparency in these fields.

While in Political Science there appear to have been no instances prior to the recent LaCour scandal[19] of a journal, university, press, or other third-party publicly and definitively establishing (or an author confessing to) fabrication or falsification,[20] there are no apparent reasons why research misconduct ought to be assumed to be a non-issue in our discipline. There is, likewise, no cause to think that Political Science is any less afflicted by "lesser" forms of questionable research practice, including, most significantly, the selective interpretation and reporting of evidence and results in favor of stronger or more striking findings. Transparency can help research communities identify and minimize such problematic practices, as well as entirely unintentional or unconscious sources of bias or error.

A related, important motivation for the push to establish transparency rules (and to make compliance a prerequisite for publication) in Political Science is the widespread sense that there are serious shortcomings in current practices of empirical documentation. For example, despite some long-standing appeals to provide comprehensive information about analytical procedures and access to the empirical information used,[21] the availability of replication datasets and code for quantitative work remains poor where it is not a meaningfully enforced requirement for publication.[22] And although published replication studies are very rare (even in other disciplines with a more uniformly neopositivist orientation[23]), scholars who seek to reproduce quantitative results presented in published studies routinely fail, even when using data and/or code provided by

---

[11] Elman and Kapiszewski 2014, 44–46. For a substantive example of an especially analytically explicit form of process tracing, see the appendix to Fairfield 2013. For a quantitative representation of the analytic steps involved in process tracing, employing Bayesian logic, see Bennett 2015 and Humphreys and Jacobs forthcoming.

[12] APSA 2012, 9f; Lupia and Elman 2014, 21f; Elman and Kapieszewski 2014, 45; Moravcsik 2014.

[13] Note that judging the *representativeness* of the evidence presented in a research output might necessitate an even broader form of data access than is often demanded by journals that currently require replication data: not just access to the sources, cases, or variables used in the analyses presented, but *all* sources consulted and all cases/variables explored by the researcher. That higher standard, however, presents difficult problems of implementation and enforcement.

[14] For an overview, see LaFollette 1996. The U.S. National Science Foundation defines research misconduct as (one or more of) fabrication, falsification and plagiarism. See NSF final rule of 7 March 2002, revising its misconduct in science and engineering regulations, promulgated via the *Federal Register* vol. 67, no 5 of 18 March 2002 (45 Code of Federal Regulations, section 689.1).

[15] Fang and Casadevall 2011.

[16] Fang, Steen, and Casadevall 2012; Steen 2011. The frequency with which retractions are attributable to research misconduct have been shown to be a substantial undercount: Of 119 papers where a research ethics investigation resulted in a finding of misconduct and subsequent retraction of the publication, 70 (58.8%) of the retrac-

tions made no mention of misconduct or research ethics as a reason for the retraction (see Resnik 2015).

[17] Hoffer 2004; Wiener 2004; DeSantis 2015.

[18] Most notably S. Walter Poulshock's (1965) fabrication of large numbers of passages in his book *Two Parties and the Tariff* in the 1880s and serious irregularities in the references and descriptive statistics in Michael Bellesiles' 2001 Bancroft-prize-winning book about the origins of U.S. gun culture, *Arming America*, leading to the author's resignation from his tenured position at Emory and the revocation of the Prize. On the latter, see Columbia University 2002; Emory University 2002; Katz et al 2002.

[19] The article involved, which had appeared in *Science*, was LaCour and Green 2014; the problems were first revealed in Broockman, Kalla and Aranow 2014. See also Singal 2015.

[20] Even disclosed cases of plagiarism are rare. A clear case is the article by Papadoulis (2005), which was retracted by the editors of the *Journal of Common Market Studies* for undue similarities to Spanou (1998); see Rollo and Paterson 2005. For an account of another case, see Lanegran 2004.

[21] See, e.g., King 1995.

[22] Dafoe 2014; Lupia and Alter 2014, 54–56.

[23] Reported frequencies of publications including replication range from a low of 0.0013 (0.13%) in education research (Makel and Plucker 2014, 308) to a high of just over 0.0107 (1.07%) in psychology (Makel, Plucker, and Hegarty 2012, 538f).

the authors.[24] Unfortunately, none of these issues are limited to quantitative research. The sharing or depositing of qualitative evidence is still relatively rare—though the technological and institutional opportunities for doing so have recently expanded (as we discuss further in our concluding essay). And when scholars have on their own attempted to track down the original sources underlying another scholar's published qualitative work, they have sometimes reported difficulty in confirming that sources support the claims for which they are cited; confirming that the cited sources are a fair representation of the known, pertinent sources; or confirming the existence of the cited sources, given the information provided in the original publication.[25]

To give another example: Based on a review of public opinion research published in the top journals in Political Science and Sociology, as well as the top specialty journals, from 1998 to 2001, Smith found that only 11.5% of articles reported a response rate along with "at least some" information on how it had been calculated.[26] While here, too, substantial strides toward greater transparency have been made in recent years in the form of a refinement of the American Association for Public Opinion Research standard (and the development of a free online calculator that implements that standard),[27] technological developments in this field are ironically undercutting attempts to raise transparency, with increasingly popular online opt-in surveys lacking an agreed response metric.[28] Here again, the problem does not appear to be unique to quantitative research. Those who collect qualitative observations often do

not provide detailed information about their data-gathering procedures, either. Consider, for instance, how rarely scholars drawing on elite interviews fully explain how they selected interview subjects, what proportion of subjects agreed to be interviewed, what questions were asked, and other procedural elements that may be relevant to the interpretation of the responses and results.[29] And as Elman and Kapiszewski point out, process tracing scholars often leave opaque how exactly key pieces of evidence have been interpreted or weighed in the drawing of inferences.[30]

It is possible that some kinds of qualitative research might not be subject to certain problems that transparency is supposed to address. The convention of considering quantitative results "statistically significant" only when the estimated coefficient's p-value falls below 0.05, for instance, appears to create incentives to selectively report model specifications that yield, for the variable(s) favored by the authors, p-values below this threshold.[31] To the extent that qualitative scholarship is more acceptant of complex empirical accounts and open to the role of contingency in shaping outcomes, it should be less subject to this particular misalignment of incentives. And to the extent that qualitative researchers are expected to point to specific events, statements, sequences, and other details of process within a case, the manipulation of empirical results might require a more conscious and direct misrepresentation of the empirical record—and might be easier for other scholars with deep knowledge of the case to detect—which, in turn, might reduce the likelihood of this form of distortion in some types of case-oriented research. Also, a number of the incentive problems that transparency is intended to address are arguably primarily problems of positivist hypothesis-testing, and would operate with less force in work that is openly inductive or interpretive in character.

At a more fundamental level, however, it is likely that many of the problems that have been observed in other disciplines and in other research traditions also apply to many forms of qualitative scholarship in Political Science. Even for qualitative work, political science editors and reviewers seem to have a preference for relatively simple explanations and tidy, theory-consistent empirical patterns—and probably especially so at the most competitive and prestigious journals, as Saey reports for the Life Sciences.[32] Such journal preferences may create unwelcome incentives for case-study researchers to select and interpret evidence in ways consistent with elegant causal accounts. Moreover, the general cognitive tendency to disproportionately search for information that confirms one's prior beliefs and to interpret information in ways favorable to those beliefs—known as confirmation bias[33]—is likely to operate as powerfully for qualitative as for quantitative researchers. Finally, there is little reason to believe that qualitative scholars

---

[24] For recent findings of such problems in political science/political economy, see Bermeo 2016 and Büthe and Morgan 2015; see also Lupia and Alter 2014, 57 and the "Political Science Replication" blog. For an illustration of the difficulty of overcoming the problem, long recognized in Economics, compare Dewald, Thursby, and Anderson (1986) with McCullough, McGeary, and Harrison (2008) and Krawczyk and Reuben (2012).

[25] E.g., Lieshout, Segers, and van der Vleuten (2004) report having been unable to locate several of the sources invoked by Moravcsik (1998) in a key section of his book; from the sources that they did find, they provide numerous examples of quotes that they allege either do not support or contradict the claim for which those sources were originally cited. (Moravcsik's reply to the critique is still forthcoming.) Lieshout (2012) takes Rosato (2011) to task for "gross misrepresentations of what really happened" due to "selective quoting" from the few documents and statements from the "grabbag of history" that support his argument and "ignoring unwelcome counter-evidence" even when his list of references suggests he must have been familiar with that counterevidence. Moravcsik (2013, 774, 790), also reviewing Rosato (2011), alleges pervasive "overtly biased" use of evidence, including "a striking number of outright misquotations, in which well-known primary and secondary sources are cited to show the diametrical opposite of their unambiguous meaning [to the point that it] should disqualify this work from influencing the debate on the fundamental causes of European integration." For the author's reply, see Rosato 2013.

[26] Smith 2002, 470.

[27] See http://www.aapor.org/AAPORKentico/Education-Resources/ For-Researchers/Poll-Survey-FAQ/Response-Rates-An-Overview. aspx (last accessed 7/20/2015).

[28] Callegaro and DiSogra 2008.

[29] See Bleich and Pekkanen's essay in this symposium for a detailed discussion of these issues.

[30] Elman and Kapiszewski 2014, 44.

[31] Gerber and Malhotra 2008.

[32] Saey 2015, 24.

[33] See, e.g., Jervis 1976, esp. 128ff, 143ff, 382ff; Tetlock 2005.

amassing and assessing vast amounts of empirical information from disparate sources are any less prone to making basic, unintentional errors than are scholars working with quantitative datasets. Transparency promises to help reduce and detect these systematic and unsystematic forms of error, in addition to facilitating understanding, interpretation, assessment, and learning.

The Symposium

Advocates of greater research transparency have brought important issues to the top of the disciplinary agenda, opening a timely conversation about the meaning, benefits, and limitations of research transparency in political inquiry. Among the central challenges confronting the push for greater transparency is the tremendous diversity in research methodologies and intellectual orientations within the discipline, even just among scholars using "qualitative" methods of empirical inquiry.

This symposium is an effort to engage scholars from a wide variety of qualitative and multi-method research traditions in a broad examination of the meaning of transparency, in line with Elman and Kapiszewski's call for "scholars from diverse research communities [to] participate and … identify the levels and types of transparency with which they are comfortable and that are consistent with their modes of analysis."[34] As a starting point for analysis, we asked contributors to this symposium to address the following questions:

- What do scholars in your research tradition most need to be transparent about? How can they best convey the most pertinent information about the procedures used to collect or generate empirical information ("data") and/or the analytic procedures used to draw conclusions from the data?

- For the type of research that you do, what materials can and should be made available (to publication outlets and ultimately to readers)?

- For the kind of work that you do, what do you see as the benefits or goals of research transparency? Is potential replicability an important goal? Is transparency primarily a means of allowing readers to understand and assess the work? Are there other benefits?

- What are the biggest challenges to realizing transparency and data access for the kind of work that you do? Are there important drawbacks or limitations?

We also asked authors to provide specific examples that illustrate the possibilities of achieving transparency and key challenges or limitations, inviting them to draw on their own work as well as other examples from their respective research traditions.

The resulting symposium presents a broad-ranging conversation about how to conceptualize and operationalize re-

search transparency across diverse forms of scholarship. The essays that follow consider, in part, how the principles of openness embedded in the APSA *Ethics Guide* and the DA-RT Statement should be put into practice in specific research traditions. Yet, the symposium also explores the appropriateness of these standards for diverse forms of scholarship. Some authors examine, for instance, the ontological and epistemological assumptions implicit in a conceptualization of empirical inquiry as the extraction of information from the social world.[35] Others make the case for an even more ambitious transparency agenda: one that attends, for instance, to the role of researcher positionality and subjectivity in shaping the research process. Contributors, further, grapple with the intellectual and ethical tradeoffs involved in the pursuit of transparency, especially as they arise for case-study and field researchers. They highlight possible costs ranging from losses in the interpretibility of findings to physical risks to human subjects.

We have organized the contributions that follow into three broad sections. The first is comprised of two essays that address the application of transparency standards to the use of common *forms of qualitative evidence*. Erik Bleich and Robert Pekkanen consider transparency in the use of interview data, while Marc Trachtenberg focuses on issues of transparency that arise particularly prominently when using primary and secondary written sources.

The articles in the second group examine research transparency in the context of differing *field-research approaches and settings*. Katherine Cramer investigates the meaning of transparency in the ethnographic study of public opinion; Victor Shih explores the distinctive challenges of openness confronting fieldwork in repressive, non-democratic settings; Sarah Elizabeth Parkinson and Elisabeth Jean Wood discuss the demands and dilemmas of transparency facing researchers working in contexts of political violence; and Timothy Pachirat interrogates the meaning of transparency from the general perspective of interpretive ethnography.

Our third cluster of essays is organized around specific *analytic methods*. David Romney, Brandon Stewart, and Dustin Tingley seek to operationalize DA-RT principles for computer-assisted text analysis; Claudius Wagemann and Carsten Schneider discuss transparency guidelines for Qualitative Comparative Analysis; Andrew Davison probes the meaning of transparency for hermeneutics; and Tasha Fairfield considers approaches to making the logic of process tracing more analytically transparent.

In our concluding essay, we seek to make sense of the variety of perspectives presented across the ten essays—mapping out considerable common ground as well as key disagreements over the meaning of research transparency and the appropriate means of achieving it. We also contemplate potential ways forward for the pursuit of greater research transparency in political science. We see the push for greater transparency as a very important development within our profession— with tremendous potential to enhance the integrity and public relevance of political science scholarship. Yet, we are also struck by—and have learned much from our contributors about—the

---

[34] Elman and Kapiszewski 2014, 46.
[35] See Lupia and Elman 2014, 20.

serious tensions and dilemmas that this endeavor confronts. In our concluding piece, we thus consider a number of ways in which the causes of transparency, intellectual pluralism, and ethical research practice might be jointly advanced. These include the development of explicitly differentiated openness standards, adjustments to DA-RT language, and reforms to graduate education and editorial practices that would broadly advance the cause of scholarly integrity.

References

APSA, Committee on Professional Ethics, Rights and Freedoms. 2012. *A Guide to Professional Ethics in Political Science. Second Edition*. Washington, DC: APSA. (http://www.apsanet.org/portals/54/Files/Publications/APSAEthicsGuide 2012.pdf, last accessed 7/16/2015)

Bennett, Andrew. 2015. "Appendix: Disciplining Our Conjectures: Systematizing Process Tracing with Bayesian Analysis." In *Process Tracing in the Social Sciences: From Metaphor to Analytic Tool*, edited by Andrew Bennett and Jeffrey Checkel. New York: Cambridge University Press: 276–298.

Bermeo, Sarah B. 2016. "Aid Is Not Oil: Donor Utility, Heterogenous Aid, and the Aid-Democracy Relationship." *International Organization* vol. 70 (forthcoming).

Broockman, David, Joshua Kalla, and Peter Aronow. 2015. "Irregularities in LaCour (2014)." Report; online at http://stanford.edu~dbroock/broockman_kalla_aronow_lg_irregularities.pdf (last accessed 7/15/2015).

Büthe, Tim, and Stephen N. Morgan. 2015. "Antitrust Enforcement and Foreign Competition: Special Interest Theory Reconsidered." (Paper presented at the Annual Meeting of the Agricultural and Applied Economics Association, San Francisco, 26-28 July 2015).

Callegaro, Mario, and Charles DiSogra. 2008. "Computing Response Metrics for Online Panels." *Public Opinion Quarterly* vol. 72, no. 5: 1008–1032.

Columbia University, Office of Public Affairs. 2002. "Columbia's Board of Trustees Votes to Rescind the 2001 Bancroft Prize." 16 December 2002. Online at http://www.columbia.edu/cu/news/02/12/bancroft_prize.html (last accessed 7/7/2015).

Dafoe, Alan. 2014. "Science Deserves Better: The Imperative to Share Complete Replication Files." *PS: Political Science and Politics* vol.47, no.1: 60–66.

"Data Access and Research Transparency (DA-RT): A Joint Statement by Political Science Journal Editors." 2014. At http://media.wix.com/ugd/fa8393_da017d3fed824cf587932534c860ea25.pdf (last accessed 7/10/2015).

DeSantis, Nick. 2015. "Arizona State Professor is Demoted After Plagiarism Inquiry." *The Chronicle of Higher Education* online, 13 July. http://chronicle.com/blogs/ticker/arizona-state-professor-demoted-after-plagiarism-inquiry/ (last accessed 7/14/2015).

Dewald, William G., Jerry G. Thursby, and Richard G. Anderson. 1986. "Replication in Empirical Economics: The Journal of Money, Credit and Banking Project." *American Economic Review* vol. 76, no. 4: 587–603.

Elman, Colin, and Diana Kapiszewski. 2014. "Data Access and Research Transparency in the Qualitative Tradition." *PS: Political Science and Politics* vol. 47, no. 1: 43–47.

Emory University, Office of Public Affairs. 2002. "Oct.25: Michael Bellesisles Resigns from Emory Faculty." Online at http://www.emory.edu/news/Releases/bellesiles1035563546.html (last accessed 7/7/2015).

Fairfield, Tasha. 2013. "Going Where the Money Is: Strategies for Taxing Economic Elites in Unequal Democracies." *World Development* vol.47 (July): 43–47.

Fang, Ferric C., and Arturo Casadevall. 2011. "Retracted Science and the Retraction Index." *Infection and Immunity* vol. 79, no. 10: 3855–3859.

Fang, Ferric C., R. Grant Steen, and Arturo Casadevall. 2012. "Misconduct Accounts for the Majority of Retracted Scientific Publications." *Proceedings of the National Academy of Sciences of the United States* vol. 109, no. 42: 17028–17033.

Gerber, Alan S., and Neil Malhotra. 2008. "Do Statistical Reporting Standards Affect What Is Published? Publication Bias in Two Leading Political Science Journals." *Quarterly Journal of Political Science* vol. 3, no.3: 313–326.

Hoffer, Peter Charles. 2004. *Past Imperfect: Facts, Fictions, Fraud—American History from Bancroft and Parkman to Ambrose, Bellesiles, Ellis and Goodwin*. New York: Public Affairs.

Humphreys, Macartan, and Alan M. Jacobs. Forthcoming. "Mixing Methods: A Bayesian Approach." *American Political Science Review*.

Jervis, Robert. 1976. *Perception and Misperception in International Politics*. Princeton: Princeton University Press.

Katz, Stanley N., Hanna H. Gray, and Laurel Thatcher Ulrich. 2002. "Report of the Investigative Committee in the Matter of Professor Michael Bellesiles." Atlanta, GA: Emory University.

King, Gary. 1995. "Replication, Replication." *PS: Political Science and Politics* vol. 28, no. 3: 444–452.

Krawczyk, Michael, and Ernesto Reuben. 2012. "(Un)available Upon Request: Field Experiment on Researchers' Willingness to Share Supplemental Materials." *Accountability in Research: Policies & Quality Assurance* vol. 19, no. 3: 175–186.

LaCour, Michael J., and Donald P. Green. 2014. "When Contact Changes Minds: An Experiment on Transmission of Support for Gay Equality." *Science* vol. 346, no. 6215: 1366–1369.

LaFollette, Marcel C. 1996. *Stealing into Print: Fraud, Plagiarism, and Misconduct in Scientific Publishing*. Berkeley: University of California Press.

Lanegran, Kimberly. 2004. "Fending Off a Plagiarist." *Chronicle of Higher Education* vol. 50, no. 43: http://chronicle.com/article/Fending-Off-a-Plagiarist/44680/ (last accessed 7/23/15).

Lieshout, Robert H. 2012. "Europe United: Power Politics and the Making of the European Community by Sebastian Rosato (review)." *Journal of Cold War Studies* vol.14, no.4: 234–237.

Lieshout, Robert H., Mathieu L. L. Segers, and Anna M. van der Vleuten. 2004. "De Gaulle, Moravcsik, and The Choice for Europe." *Journal of Cold War Studies* vol. 6, no. 4: 89–139.

Lupia, Arthur, and George Alter. 2014. "Openness in Political Science: Data Access and Research Transparency." *PS: Political Science and Politics* vol. 47, no. 1: 54–59.

Lupia, Arthur, and Colin Elman. 2014. "Openness in Political Science: Data Access and Research Transparency." *PS: Political Science and Politics* vol. 47, no. 1: 19–42.

Makel, Matthew C., and Jonathan A. Plucker. 2014. "Facts Are More Important Than Novelty: Replication in the Education Sciences." *Educational Researcher* vol. 43, no. 6: 304–316.

Makel, Matthew C., Jonathan A. Plucker, and Boyd Hegarty. 2012. "Replications in Psychology Research: How Often Do They Really Occur?" *Perspectives in Psychological Science* vol. 7, no. 6: 537–542.

McCullough, B. D., Kerry Anne McGeary, and Teresa D. Harrison. 2008. "Do Economics Journal Archives Promote Replicable Research?" *Canadian Journal of Economics/Revue canadienne d'économique* vol. 41, no. 4: 1406–1420.

McGuire, Kevin T. 2010. "The Was a Crooked Man(uscript): A Not-So-Serious Look at the Serious Subject of Plagiarism." *PS: Political Science and Politics* vol.43, no.1: 107–113.

Moravcsik, Andrew. 1998. *The Choice for Europe: Social Purpose and State Power from Messina to Maastricht*. Ithaca, NY: Cornell University Press.

———. 2013. "Did Power Politics Cause European Integration? Realist Theory Meets Qualitative Methods." *Security Studies* vol. 22, no. 4: 773–790.

———. 2014. "Transparency: The Revolution in Qualitative Research." *PS: Political Science and Politics* vol. 47, no. 1: 48–53.

Papadoulis, Konstantinos J. 2005. "EU Integration, Europeanization and Administrative Convergence: The Greek Case." *Journal of Common Market Studies* vol. 43, no. 2: 349–370.

"Political Science Replication." https://politicalsciencereplication.wordpress.com/ (last accessed 7/20/15).

Resnik, David B. 2015. "Scientific Retractions and Research Misconduct." (Presentation at the Science and Society Institute, Duke University, 9 April 2015).

Rollo, Jim, and William Paterson. 2005. "Retraction." *Journal of Common Market Studies* vol. 43, no. 3: i.

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

———. 2013. "Theory and Evidence in Europe United: A Response to My Critics." *Security Studies* vol. 22, no. 4: 802–820.

Saey, Tina Hesman. 2015. "Repeat Performance: Too Many Studies, When Replicated, Fail to Pass Muster." *Science News* vol. 187, no. 2: 21–26.

Singal, Jesse. 2015. "The Case of the Amazing Gay-Marriage Data: How a Graduate Student Reluctantly Uncovered a Huge Scientific Fraud." *New York Magazine* online 29 May: http://nymag.com/scienceofus/2015/05/how-a-grad-student-uncovered-a-huge-fraud.html (last accessed 7/15/2015)

Smith, Tom W. 2002. "Reporting Survey Nonresponse in Academic Journals." *International Journal of Public Opinion Research* vol. 14, no. 4: 469–474.

Spanou, Calliope. 1998. "European Integration in Administrative Terms: A Framework for Analysis and the Greek Case." *Journal of European Public Policy* vol. 5, no. 3: 467–484.

Steen, R. Grant. 2011. "Retractions in the Scientific Literature: Is the Incidence of Research Fraud Increasing?" *Journal of Medical Ethics* vol. 37, no. 4: 249–253.

Tetlock, Philip E. 2005. *Expert Political Judgment: How Good Is It? How Can We Know?* Princeton: Princeton University Press.

Wiener, Jon. 2004. *Historians in Trouble: Plagiarism, Fraud and Politcs in the Ivory Tower*. New York: The New Press.

Transparency About Qualitative Evidence

Data Access, Research Transparency, and Interviews: The Interview Methods Appendix

Erik Bleich
Middlebury College

Robert J. Pekkanen
University of Washington

Interviews provide a valuable source of evidence, but are often neglected or mistrusted because of limited data access for other scholars or inadequate transparency in research production or analysis. This incomplete transparency creates uncertainty about the data and leads to a "credibility gap" on interview data that has nothing to do with the integrity of the researcher. We argue that addressing transparency concerns head-on through the creation of common reporting standards on interview data will diminish this uncertainty, and thus benefit researchers who use interviews, as well as their readers and the scholarly enterprise as a whole. As a concrete step, we specifically advocate the adoption of an "Interview Methods Appendix" as a reporting standard. Data access can involve difficult ethical issues such as interviewee confidentiality, but we argue that the more data access interviewers can provide, the better. The guiding principles of the Statement on Data Access and Research Transparency (DA-RT) will also enhance scholarship that utilizes interviews.

A flexible and important method, interviews can be employed for preliminary research, as a central source for their study, or as one part of a multi-method research design.[1] While interviewing in the preliminary research stage can provide a very efficient way to generate research questions and hypotheses,[2] our focus here will be on how research transparency can increase the value of interviews used in the main study or as one leg of a multi-methods research design. Some types of interviews, whether interstitial or simply preliminary, are not as essential to report. However, when interviews are a core part of the research, transparency in production and analysis is critical. Although our arguments apply to a variety of sampling strategies, as discussed below, we think they are especially germane to purposive- or snowball-sampled interviews in a main study.[3]

The recent uptick in interest in qualitative methods has drawn more attention to interviews as a method. However, the benefits of emerging advances in interview methodology will only be fully realized once scholars agree on common reporting standards for data access and research transparency.

Erik Bleich is Professor of Political Science at Middlebury College. He can be found online at ebleich@middlebury.edu and http://www.middlebury.edu/academics/ps/faculty/node/25021. Robert Pekkanen is Professor at the Henry M. Jackson School of International Studies, University of Washington. He can be found online at pekkanen@ uw.edu and http://www.robertpekkanen.com/. A fuller version of many of the arguments we make here can be found in Bleich and Pekkannen (2013) in a volume entirely devoted to interview research in political science (Mosley 2013).

---

[1] Lynch 2013, esp. 34–38.

[2] Lynch 2013, 34f; Gallagher 2013, 183–185. For a more general discussion of the importance of using the empirical record to develop research questions and hypotheses, see Gerring 2007, esp. 39–43.

[3] See Lynch (2013) for a lucid discussion of random, purposive, convenience, snowball and interstitial interviewing.

Main Issues with Interviews and Proposed Solutions

The use and acceptance of interviews as evidence is limited by concerns about three distinct empirical challenges: how to define and to sample the population of relevant interviewees (sampling); whether the interviews produce valid information (validity); and whether the information gathered is reliable (reliability).[4] We argue that research transparency and data access standards can mitigate these concerns and unlock an important source of evidence for qualitative researchers. We discuss examples of an Interview Methods Appendix and an Interview Methods Table here as a way to convey substantial information that helps to overcome these concerns. At the same time, we recognize that standards for reporting information will be subject to some variation depending on the nature of the research project and will evolve over time as scholars engage with these discussions. For us, the central goal is the self-conscious development of shared standards about what "essential" information should be reported to increase transparency. Our Interview Methods Appendix and Interview Methods Table are meant as a concrete starting point for this discussion.

In this context, it is useful to consider briefly the differences between surveys and interviews. Surveys are widely deployed as evidence by scholars from a variety of disciplines. Like interviews, surveys rely on information and responses gained from human informants. There are many well-understood complications involved in gathering and assessing survey data. Survey researchers respond to these challenges by reporting their methods in a manner that enables others to judge how much faith to place in the results. We believe that if similar criteria for reporting interview data were established, then interviews would become a more widely trusted and used source of evidence. After all, surveys can be thought of as a collection of short (sometimes not so short) interviews. Surveys and interviews thus fall on a continuum, with trade-offs between large-n and small-n studies. Just as scholars stress the value of both types of studies, depending on the goal of the researchers,[5] both highly structured survey research and semi- or unstructured "small-n" interviews, such as elite interviews, should have their place in the rigorous scholar's tool kit.[6]

---

[4] Here we follow Mosley's (2013a, esp. 14–26) categorization, excluding only her normative category of the challenge of ethics, which we address only as it relates to DA-RT issues. See also Berry (2002) on issues of validity and reliability.

[5] See Lieberman 2005, 435.

[6] Many scholars believe that interviews provide a wealth of contextual data that are essential to their interpretations, understanding, or analysis. We do not argue that an Interview Methods Appendix can communicate every piece of potentially relevant information, but rather that it can convey a particular range of information in order to increase the credibility of interviews as a research method. Information not summarized in the Interview Methods Appendix may include what Mosley calls "meta-data" (2013a, 7, 22, 25). Exactly *what* information must be reported (and what can be omitted) is subject to a process of consensual scholarly development; we view our suggestions as a starting point for this process.

Qualitative social scientists can benefit from a common set of standards for reporting their data so that readers and reviewers can judge the value of their evidence. As with quantitative work, it will be impossible for qualitative researchers to achieve perfection in their methods, and interview-based work should not be held to an unrealistic standard. But producers and consumers of qualitative scholarship profit from being more conscious about the methodology of interviewing and from being explicit about reporting uncertainty. Below, we discuss how transparent reporting works for researchers engaged in purposive sampling and for those using snowballing techniques.

Sampling: Purposive Sampling Frames and Snowball Additions

It is often possible for the researcher to identify a purposive, theoretically motivated set of target interviewees prior to going into the field. We believe that doing this in advance of the interviews, and then reporting interviews successfully obtained, requests refused, and requests to which the target interviewee never responded, has many benefits. For one, this kind of self-conscious attention to the sampling frame will allow researchers to hone their research designs before they enter the field. After identifying the relevant population of actors involved in a process, researchers can focus on different classes of actors within that population—such as politicians, their aides, civil servants from the relevant bureaucracies, NGOs, knowledgeable scholars and journalists, etc.—different types within the classes—progressive and conservative politicians, umbrella and activist NGOs, etc.—and/or actors involved in different key time periods in a historical process—key players in the 1980, 1992, and 2000 elections, etc. Drawing on all classes and types of actors relevant to the research project helps ensure that researchers receive balanced information from a wide variety of perspectives. When researchers populate a sampling frame from a list created by others, the source should be reported—whether that list is of sitting parliamentarians or business leaders (perhaps drawn from a professional association membership) or some other roster of a given population of individuals.

Often used as a supplement to purposive sampling, "snowball" sampling refers to the process of seeking additional contacts from one's interviewees. For populations from which interviews can be hard to elicit (say, U.S. Senators), the technique has an obvious attraction. Important actors approached with a referral in hand are more likely to agree to an interview request than those targeted through "cold-calls." In addition, if the original interviewee is a good source, then she is likely to refer the researcher to another knowledgeable person. This snowball sampling technique can effectively reveal networks or key actors previously unknown to the researcher, thereby expanding the sampling frame. All in all, snowballing has much to commend it as a technique, and we do not argue against its use. However, we do contend that the researcher should report interviewee-list expansion to readers and adjust the sampling frame accordingly if necessary to maintain a balanced set

of sources.

Non-Response Rates and Saturation

Reporting the sampling frame is a vital first step, but it is equally important to report the number of interviews sought within the sampling frame, the number obtained, and the number declined or unavailable. Drawing a parallel with large-n studies, survey researchers are generally advised to report response rates, because the higher the response rate, the more valid the survey results are generally perceived to be. Such non-response bias might also skew results in the interviewing process. In a set of interviews about attitudes towards the government, for example, those who decline to participate might do so because of a trait that would lead them to give a particular type of response to the interviewer's questions, either systematically positive or negative. If so, then we would be drawing inferences from our conducted interviews that would be inaccurate, because we would be excluding a set of interviewees that mattered a great deal to the validity of our findings. Under current reporting practices, we have no way to assess response rates or possible non-response bias in interviews. At present, the standard process involves reporting who was interviewed (and some of what was said), but not whom the author failed to reach, or who declined an interview.

In addition, to allow the reader to gauge the validity of the inferences drawn from interviews, it is crucial for the researcher to report whether she has reached the point of saturation. At saturation, each new interview within and across networks reveals no new information about a political or policymaking process.[7] If respondents are describing the same causal process as previous interviewees, if there is agreement across networks (or predictable disagreement), and if their recommendations for further interviewees mirror the list of people the researcher has already interviewed, then researchers have reached the point of saturation. Researchers must report whether they reached saturation to help convey to readers the relative certainty or uncertainty of any inferences drawn from the interviews.

In practice, this may involve framing the interview reporting in a number of different ways. For simplicity's sake, our hypothetical Interview Methods Table below shows how to report saturation within a set of purposively sampled interviewees. However, it may be more meaningful to report saturation with respect to a particular micro- or meso-level research question. To give an example related to the current research of one of us, the interview methods appendix may be organized around questions such as "Do the Left-Right political ideologies of judges affect their rulings in hate speech cases?" (high saturation), and "Were judges' hate speech decisions motivated by sentiments that racism was a pressing social problem in the period 1972-1983?" (low saturation). Reporting low saturation does not mean that research inferences drawn from the responses are invalid, only that the uncertainty around those inferences is higher, and that the author could increase confi-

dence in her conclusions by seeking information from non-interview-based sources. Reporting levels of confidence in and uncertainty about findings are critical to research transparency.

Data Access

Once researchers have conveyed to readers that they have drawn on a valid sample, they face the task of convincing observers that the inferences drawn from those responses are similar to those that would be drawn by other researchers looking at the same data. In many ways, the ideal solution to this dilemma is to post full interview transcripts on a web site so that the curious and the intrepid can verify the data themselves. This standard of qualitative data archiving should be the discipline's goal, and we are not alone in arguing that it should move in this direction.[8] At the same time, we fully recognize that it will be impractical and even impossible in many cases. Even setting aside resource constraints, interviews are often granted based on assurances of confidentiality or are subject to conditions imposed by human subject research, raising not only practical, but also legal and ethical issues.[9]

Whether it is possible to provide full transcripts, redacted summaries of interviews, or no direct information at all due to ethical constraints, we think it is vital for researchers to communicate the accuracy of reported interview data in a rigorous manner. In many scenarios, the researcher aims to convey that the vast majority of interviewees agree on a particular point. Environmental lobbyists may judge a conservative government unsympathetic to their aims, or actors from across the political spectrum may agree on the importance of civil society groups in contributing to neighborhood policing. Rather than simply reporting this general and vague sentiment, in most instances it is possible to summarize the number of lobbyists expressing this position as a percentage of the total lobbyists interviewed and as a percentage of the lobbyists specifically asked or who spontaneously volunteered their opinion on the government's policy. Similarly, how many policymakers and politicians were interviewed, and what percentage expressed their enthusiasm for civil society groups? This is easiest to convey if the researcher has gone through the process of coding interviews that is common in some fields. It is more difficult to assess if scholars have not systematically coded their interviews, but in these circumstances it is all the more important to convey a sense of the representativeness and reliability of the information cited or quoted. Alternatively, if the researcher's analysis relies heavily on information provided in one or two interviews, it is incumbent upon the author to explain the basis upon which she trusts those sources more than others. Perhaps an interviewee has provided information that runs counter to his interests and is thus judged more likely to be truthful than his counterparts, or an actor in a critical historical juncture was at the center of a network of policymakers while others were more peripheral.

---

[7] Guest, Bunce, and Johnson 2006.

[8] See Elman, Kapiszewski, and Vinuela 2010; Moravcsik 2010, 31.

[9] Parry and Mauthner 2004; Brooks 2013; MacLean 2013. See also Mosley 2013a, 14–18.

Beyond reporting the basis for trusting some interviews over others, it is useful to remember that very few studies rely exclusively on interview data for their conclusions. While other types of sources have their own weaknesses, when interview evidence is ambiguous or not dispositive, scholars can fruitfully triangulate with other sources to resolve ambiguities in the record in order to gauge and to confirm the reliability and validity of the information gathered. Perhaps no method of summary reporting will fully convince skeptics about the accuracy of information gathered through interviews. But strategies such as those suggested here will make even-handed readers more certain about the reliability of the inferences when judging the rigor of the scholarship and the persuasiveness of the argument.

To the extent that researchers are able to provide transcripts of their interviews in online appendices or qualitative data archives—perhaps following an initial embargo period standard among quantitative researchers for newly-developed datasets, or for archivists protecting sensitive personal information—there are potentially exponential gains to be made to the research community as a whole.[10] Not only will this practice assure readers that information sought, obtained, and reported accurately conveys the reality of the political or policy-making process in question, but it will also allow researchers in years to come access to the reflections of key practitioners in their own words, which would otherwise be lost to history. Imagine if in forty years a scholar could reexamine a pressing question not only in light of the written historical record, but also with your unique interview transcripts at hand. Carefully documenting interviewing processes and evidence will enhance our confidence that we truly understand political events in the present day and for decades to come.

How and What to Report: The Interview Methods Appendix and the Interview Methods Table

How can researchers quickly and efficiently communicate that they have done their utmost to engage in methodologically rigorous interviewing techniques? We propose the inclusion of an "Interview Methods Appendix" in any significant research product that relies heavily on interviews. The Interview Methods Appendix can contain a brief discussion of key methodological issues, such as: how the sample frame was constructed; response rate to interview requests and type of interview conducted (in person, phone, email, etc.); additional and snowball interviews that go beyond the initial sample frame; level of saturation among interview categories or research questions; format and length of interview (structured, semi-structured, etc.); recording method; response rates and consistency of reported opinions; and, confidence levels and compensation strategies.[11]

It is possible to include a brief discussion of these issues in an appendix of a book with portions summarized in the general methodology section, or as an online, hyperlinked addendum to an article where space constraints are typically more severe.[12] In addition, large amounts of relevant methodological information can be summarized in an Interview Methods Table. We provide an example of a table here to demonstrate its usefulness in communicating core information. This table is based on hypothetical research into the German state's management of far right political parties in the early 2000s, and thus conveys a purposive sample among non-far right politicians, far right politicians, constitutional court judges, German state bureaucrats, and anti-racist NGOs. Setting up a table in this way allows readers to understand key elements related to

Table 1: Hypothetical Interview Methods Table

Interviewee	Status	Source	Saturation	Format	Length	Recording	Transcript
Category 1			Yes				
CDU politician	Conducted in person 4/22/2004	Sample frame		Semi-structured	45 mins	Concurrent notes & supplementary notes w/i 1 hr	Confidentiality requested
SPD politician Hart	Conducted in person 4/22/2004	Sample frame & referred by CDU politician		Semi-structured	1 hr	Audio recording	transcript posted
Green politician	Conducted in person 4/23/2004	Sample frame		Semi-structured	45 mins	Concurrent notes & supplementary notes w/i 1 hr	Confidentiality requested
FDP politician Weiss	Refused 2/18/2004	Sample frame					
Die Linke politician	No response	Sample frame					
SPD politician's aide	Conducted in person 4/26/2004	Referred by SPD politician Hart		Semi-structured	1 hr 15 mins	Audio recording	Confidentiality required

---

[10] Elman, Kapiszewski, and Vinuela 2010.

[11] Bleich and Pekkanen 2013, esp. 95–105.

[12] Moravcsik 2010, 31f.

Table 1 (cont.): Hypothetical Interview Methods Table

Interviewee	Status	Source	Saturation	Format	Length	Recording	Transcript
Category 2			No				
REP politician	No response	Sample frame					
DVU politician	No response	Sample frame					
NPD politician	Accepted 3/16/2004; then declined 4/20/2004	Sample frame					
NPD lawyer	Declined 4/20/2004	Sample frame					
Category 3			No				
Constitutional Court judge 1	No response	Sample frame					
Constitutional Court judge 2	No response	Substitute in sample frame					
Category 4			Yes				
Interior Ministry bureaucrat 1	Conducted in person 4/24/2004	Sample frame		Semi-structured	45 mins	Concurrent notes & supplementary notes w/i 1 hr	Confidentiality required
Interior Ministry bureaucrat 2	Conducted in person 4/24/2004	Sample frame		Semi-structured	45 mins	Concurrent notes & supplementary notes w/i 1 hr	Confidentiality required
Justice Ministry bureaucrat	Conducted via email 4/30/2004	Referred by Interior Ministry bureaucrat 2		Structured	N/A	Email transcript	Confidentiality required
Category 5			Yes				
Anti-fascist NGO leader Korn	Conducted in person 4/22/2004	Sample frame		Semi-structured	1 hr 10 mins	Audio recording	Redacted transcript posted
Anti-fascist NGO leader Knoblauch	Conducted in person 4/25/2004	Sample frame		Semi-structured	50 mins	Audio recording	Redacted transcript posted
Anti-fascist NGO leader 3	Not sought	Referred by Anti-Fascist NGO leader Korn					
Anti-discrimination NGO leader Spitz	Conducted in person 4/29/2004	Referred by Anti-Fascist NGO leader Korn and by Far Right scholar Meyer		Semi-structured	1 hr 30 mins	Audio recording	Transcript posted
Overall			High				See www.bleichpekkanen.transcripts[*]

[*] Sample Methods Table for a hypothetical research project (see text). All names and the URL are imaginary.

sampling and to the validity and reliability of inferences. It therefore conveys the comprehensiveness of the inquiry and the confidence level related to multiple aspects of the interview component of any given research project.

We recognize that legitimate constraints imposed by Institutional Review Boards, by informants themselves, or by professional ethics may force researchers to keep some details of the interview confidential and anonymous. In certain cases, the Interview Methods Table might contain "confidentiality requested" and "confidentiality required" for every single interview. We do not seek to change prevailing practices that serve to protect informants. However, we believe that even in such circumstances, the interviewer can safely report many elements in an Interview Methods Appendix and Interview Methods Table—to the benefit of researcher and reader alike. A consistent set of expectations for reporting will give readers more confidence in research based on interview data, which in turn will liberate researchers to employ this methodology more often and with more rigor.

References

Berry, Jeffrey M. 2002. "Validity and Reliability Issues In Elite Interviewing." *PS: Political Science & Politics* vol. 35, no. 4: 679–682.

Bleich, Erik, and Robert Pekkanen. 2013. "How to Report Interview Data: The Interview Methods Appendix." In *Interview Research in Political Science*, edited by Layna Mosley. Ithaca: Cornell University Press, 84–105.

Brooks, Sarah M. 2013. "The Ethical Treatment of Human Subjects and the Institutional Review Board Process." In *Interview Research in Political Science*, edited by Layna Mosley. Ithaca: Cornell University Press, 45–66.

Elman, Colin, Diana Kapiszewski, and Lorena Vinuela. 2010. "Qualitative Data Archiving: Rewards and Challenges." *PS: Political Science & Politics* vol. 43, no. 1: 23–27.

Gerring, John. 2007. *Case Study Research: Principles and Practice*. 2nd ed. Cambridge: Cambridge University Press.

Guest, Greg, Arwen Bunce, and Laura Johnson. 2006. "How Many Interviews Are Enough? An Experiment with Data Saturation and Variability." *Field Methods* vol. 18, no. 1: 59–82.

Lieberman, Evan S. 2005. "Nested Analysis as a Mixed-Method Strategy for Comparative Research." *American Political Science Review* vol. 99, no 3: 435–452.

Lynch, Julia F. 2013. "Aligning Sampling Strategies with Analytical Goals." In *Interview Research in Political Science*, edited by Layna Mosley. Ithaca: Cornell University Press, 31–44.

MacLean, Lauren M. "The Power of the Interviewer." In *Interview Research in Political Science*, edited by Layna Mosley. Ithaca: Cornell University Press, 67–83.

Moravcsik, Andrew. 2010. "Active Citation: A Precondition for Replicable Qualitative Research." *PS: Political Science & Politics* vol. 43, no. 1: 29–35.

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

Mosley, Layna. 2013a. "Introduction." In *Interview Research in Political Science*, edited by Layna Mosley. Ithaca: Cornell University Press: 1–28.

Parry, Odette, and Natasha S. Mauthner. 2004. "Whose Data Are They Anyway?" *Sociology* vol. 38, no. 1: 132–152.

Transparency in Practice: Using Written Sources

Marc Trachtenberg
University of California, Los Angeles

Individual researchers, according to the revised set of guidelines adopted by the American Political Science Association two years ago, "have an ethical obligation to facilitate the evaluation of their evidence-based knowledge claims" not just by providing access to their data, but also by explaining how they assembled that data and how they drew "analytic conclusions" from it.[1] The assumption was that research transparency is of fundamental importance for the discipline as a whole, and that by holding the bar higher in this area, the rigor and richness of scholarly work in political science could be substantially improved.[2]

Few would argue with the point that transparency is in principle a good idea. But various problems arise when one tries to figure out what all this means in practice. I would like to discuss some of them here and present some modest proposals about what might be done in this area. While the issues that I raise here have broad implications for transparency in political research, I will be concerned here mainly with the use of a particular form of evidence: primary and secondary written sources.

Let me begin by talking about the first of the three points in the APSA guidelines, the one having to do with access to data. The basic notion here is that scholars should provide clear references to the sources they use to support their claims—and that it should be easy for anyone who wants to check those claims to find the sources in question. Of the three points, this strikes me as the least problematic. There's a real problem here that needs to be addressed, and there are some simple measures we can take to deal with it. So if it were up to me this would be the first thing I would focus on.

What should be done in this area? One of the first things I was struck by when I started reading the political science literature is the way a scholar would back up a point by citing, in parentheses in the text, a long list of books and articles, without including particular page numbers in those texts that a reader could go to see whether they provided real support for the point in question. Historians like me didn't do this kind of thing, and this practice struck me as rather bizarre. Did those authors really expect their readers to plow through those books and articles in their entirety in the hope of finding the particular passages that related to the specific claims being made? Obviously not. It seemed that the real goal was to establish the author's scholarly credentials by providing such a list. The

Marc Trachtenberg is Research Professor of Political Science at the University of California, Los Angeles. He is online at trachten@polisci.ucla.edu and http://www.sscnet.ucla.edu/polisci/faculty/trachtenberg.

[1] American Political Science Association 2012.
[2] See especially Moravcsik 2014.

whole practice did not facilitate the checking of sources; in fact, the inclusion of so much material would deter most readers from even bothering to check sources. It was amazing to me that editors would tolerate, and perhaps even encourage, this practice. But it is not unreasonable today to ask them to insist on precise page-specific references when such references are appropriate. The more general principle here is that citing sources should not be viewed as a way for an author to flex his or her academic muscles; the basic aim should be to allow readers to see, with a minimum of trouble on their part, what sort of basis there is for the claim being made. This is something journal editors should insist on: the whole process should be a lot more reader-friendly than it presently is.

A second easily-remediable problem has to do with the "scientific" system of citation that journals like the *American Political Science Review* use. With this system, references are given in parentheses in the text; those references break the flow of the text and make it harder to read. This problem is particularly serious when primary, and especially archival, sources are cited. The fact that this method makes the text less comprehensible, however, was no problem for those who adopted this system: the goal was not to make the argument as easy to understand as possible, but rather to mimic the style of the hard sciences. (When the APSR switched to the new system in June 1978, it noted that that system was the one "used by most scientific journals."[3]) It was obviously more important to appear "scientific" than to make sure that the text was as clear as possible. One suspects, in fact, that the assumption is that real science *should* be hard to understand, and thus that a degree of incomprehensibility is a desirable badge of "scientific" status. Such attitudes are very hard to change, but it is not inconceivable that journal editors who believe in transparency would at least permit authors to use the traditional system of citing sources in footnotes. It seems, in fact, that some journals in our field do allow authors to use the traditional system for that very reason.

The third thing that editors should insist on is that citations include whatever information is needed to allow a reader to find a source without too much difficulty. With archival material especially, the references given are often absurdly inadequate. One scholar, for example, gave the following as the source for a document he was paraphrasing: "Minutes of the Committee of Three, 6 November 1945, NARA, RG 59."[4] I remember thinking: "try going to the National Archives and putting in a call slip for that!" To say that this particular document was in RG59, the Record Group for the records of the State Department, at NARA—the National Archives and Records Administration—was not very helpful. RG59, as anyone who has worked in that source knows, is absolutely enormous. To get the Archives to pull the box with this document in it, you need to provide more detailed information. You need to tell them which collection within RG59 it's in, and you need to give them information that would allow them to pull the right box in that collection. As it turns out, this particular document is in the State Department Central (or Decimal) Files, and if you gave them the decimal citation for this document—which in this case happens to be 740.00119 EW/11-645—they would know which box to pull. How was I able to find that information? Just by luck: at one point, this scholar had provided a two-sentence quotation from this document; I searched for one of those sentences in the online version of the State Department's *Foreign Relations of the United States* series, and sure enough, an extract from that document containing those sentences had been published in that collection and was available online. That extract gave the full archival reference. But a reader shouldn't have to go through that much trouble to find the source for a claim.

How would it be possible to get scholars to respect rules of this sort? To a certain extent, simply pointing out, in methodological discussions like the one we're having now, how important this kind of thing is might have a positive effect, especially if people reading what we're saying here are convinced that these rules make sense. They might then make sure their students cite sources the right way. And beyond that people who do not respect those rules can be held accountable in book reviews, online roundtables, and so on. That's how norms of this sort tend to take hold.

Finally, let me note a fourth change that can be readily made. Scholars often set up their analyses by talking about what other scholars have claimed. But it is quite common to find people attributing views to other scholars that go well beyond what those scholars have actually said. Scholars, in fact, often complain about how other people have mis-paraphrased their arguments. It seems to me that we could deal with this problem quite easily by having a norm to the effect that whenever someone else's argument is being paraphrased, quotations should be provided showing that that scholar has actually argued along those lines. This would go a long way, I think, toward minimizing this problem.

Those are the sorts of changes that can be made using traditional methods. But it is important to note that this new concern with transparency is rooted, in large measure, in an appreciation for the kinds of things that are now possible as a result of the dramatic changes in information technology that have taken place over the past twenty-five years or so. All kinds of sources—both secondary and primary sources—are now readily available online, and can be linked directly to references in footnotes. When this is done, anyone who wants to check a source need only click a link in an electronic version of a book or article to see the actual source being cited. In about ten or twenty years, I imagine, we will all be required to provide electronic versions of things we publish, with hypertext links to the sources we cite. A number of us have already begun to

---

[3] Instructions to Contributors, *American Political Science Review* vol.72 no.2 (June 1978), 398 (http://www.jstor.org/stable/1954099, last accessed 6/27/2015).

[4] The citation appeared in notes 8, 10, 17, and 19 on pp. 16-18 of Eduard Mark's comment in the H-Diplo roundtable on Trachtenberg 2008 (roundtable: https://h-diplo.org/roundtables/PDF/Roundtable-X-12.pdf; original article: http://dx.doi.org/10.1162/jcws.2008.10.4.94). For my comment on Dr. Mark's use of that document, see pp. 52–54 in the roundtable.

move in that direction: I myself now regularly post on my website electronic versions of articles I write with direct links to the sources I have cited in the footnotes.[5] For the scholarly community as a whole, perhaps the most important thing here is to make sure that we have a single, unified set of standards that would govern how we adjust to, and take advantage of, the digital revolution. A possible next step for the Data Access and Research Transparency project would be to draft guidelines for book publishers and journal editors that might give them some sense for how they should proceed so that whatever norms do emerge do not take shape in a purely haphazard way.

But what about the two other forms of transparency called for in the guidelines? Individual researchers, the APSA *Guide* says, "should offer a full account of the procedures used to collect or generate" their data, and they "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." What are we to make of those precepts?

Let's begin with the first one—with what the guidelines call "production transparency," the idea that researchers "should offer a full account of the procedures used to collect or generate the data." The goal here was to try to counteract the bias that results from people's tendency to give greater weight to evidence that supports their argument than to that which does not.[6] And it is certainly true that this problem of "cherry-picking," as it is called, deserves to be taken very seriously. But I doubt whether this guideline is an effective way of dealing with it. People will always say that their sources were selected in an academically respectable way, no matter how good, or how bad, the process really is. Forcing people to explain in detail how they have collected their data will, I'm afraid, do little to improve the situation.

To show what I mean, let me talk about an article I did a few years ago dealing with audience costs theory—that is, with the claim that the ability of a government to create a situation in which it would pay a big domestic political price for backing down in a crisis plays a key role in determining how international crises run their course.[7] I identified a whole series of crises in which one might expect the audience costs mechanism to have played a role. I then looked at some historical sources relating to each of those cases to see whether that mechanism did in fact play a significant role in that particular case; my conclusion was that it did not play a major role in any of those cases. If I had been asked to explain how I collected my data, I would have said "I looked at all the major historical sources—important books and articles plus easily available collections of diplomatic documents—to see what I could find

that related to the issue at hand." That I was using that method should already have been clear to the reader—all he or she would have had to do was look at my footnotes—so including an explanation of that sort would contribute little. But maybe I would be expected to go further and explain in detail the process I used to decide which sources were important and which were not. What would this entail? I could explain that I searched in JSTOR for a number of specific search terms; that after identifying articles that appeared in what I knew were well-regarded journals, I searched for those articles in the Web of Science to see how often they were cited and to identify yet further articles; that I supplemented this by looking at various bibliographies which I identified in various ways; and that having identified a large set of studies, I then looked at them with the aim of seeing how good and how useful they were for my purposes. I could go further still and present a detailed analysis of the intellectual quality of all those secondary works, including works I had decided not to use. That sort of analysis would have been very long—much longer, in fact, than the article itself—but even if I did it, what would it prove? The works I liked, a reader might well think, were the works I happened to find congenial, because they supported the argument I was making. And even if the reader agreed that the studies I selected were the best scholarly works in the area, how could he or she possibly know that I had taken into account all of the relevant information found in that material, regardless of whether it supported my argument? I have my own ways of judging scholarly work, and I doubt whether my judgment would be much improved if authors were required to lay out their methods in great detail.

I also wonder about how useful the third guideline, about "analytic transparency," will be in practice. The idea here is that researchers "should provide a full account of how they draw their analytic conclusions from the data"—that is, that they should "clearly explicate the links connecting data to conclusions." It sometimes seems that the authors of the guidelines think they are asking scholars to do something new—to provide a statement about method that would be a kind of supplement to the book or article in question. But my basic feeling is that scholarly work, if it is any good at all, should *already* do this. And it is not just that a scholarly work should "explicate the links connecting data to conclusions," as though it is just one of a number of things that it should do. My main point here is that this is what a scholarly work *is*, or at least what it should be. The whole aim of a scholarly article, for example, should be to show how conclusions are drawn from the evidence.

It is for that reason that a work of scholarship should have a certain formal quality. The goal is *not* to describe the actual process (involving both research and thinking) that led to a set of conclusions. It is instead to develop an argument (necessarily drawing on a set of assumptions of a theoretical nature) that shows how those conclusions follow from, or at least are supported by, a certain body of evidence. It is certainly possible to explain in detail how in practice one reached those conclusions—I spend a whole chapter in my methods

---

[5] I have now posted five such articles: Trachtenberg 2005; 2011; 2013a; 2013b; and 2013c. I generally include a note in the published version giving the location of the electronic version and noting that it contains links to the materials cited. The texts themselves, in both versions, are linked to the corresponding listing in my c.v. (http://www.sscnet.ucla.edu/polisci/faculty/trachtenberg/cv/cv.html).

[6] Moravcsik 2014, 49.

[7] Trachtenberg 2012.

book showing how in practice one does this kind of work—but normally this is not what a work of scholarship does.[8] It is normally much leaner, and has a very different, and more formal, structure. One never talks about everything one has looked at; one tries instead to distill off the essence of what one has found and to present it as a lean and well-reasoned argument. Nine-tenths of the iceberg will—and indeed should—lie below the surface; it is important to avoid clutter and to make sure that the argument is developed clearly and systematically; the logic of the argument should be as tight as possible.

So a lot of what one does when one is analyzing a problem is bound to be missing from the final text, and it is quite proper that it should not be included. Let me again use that audience costs paper as an example. After it came out, *Security Studies* did a forum on it, and one of the criticisms that was made there had to do with what could be inferred from the historical evidence. The fact that one could not find much evidence that political leaders in a particular crisis deliberately exploited the audience costs mechanism, Erik Gartzke and Yonatan Lupu argued, does not prove much, because there is no reason to suppose that intentions would be revealed by the documentary record; absence of evidence is not evidence of absence.[9]

The answer here is that one can in certain circumstances infer a great deal from the "absence of evidence." What one can infer depends on the kind of material we now have access to—on the quality of the documentation, on what it shows about how freely political leaders express themselves when they were talking about these issues in private to each other at the time. One can often reach certain conclusions about what their intentions were in choosing to react the way they did on the basis of that material. Those conclusions might not be absolutely rock-solid—one never knows for sure what is in other people's minds—but one can often say more than zero, and sometimes a lot more than zero, about these questions. But should that point have been explained in the original paper? It would be okay to deal with it if one was writing a piece on historical methodology, but these methodological points should not have to be explained every time one is dealing with a substantive issue. For it is important to realize that you always pay a price in terms of loss of focus for dealing with ancillary matters and drifting away from the issue that is the real focus of the analysis.

My point here is that a good deal more goes into an assessment of what one is to make of the evidence than can be summed up in the sort of statement this third guideline seems to call for. Philosophers of science point out that in reaching conclusions "logical rules" are less important than "the mature sensibility of the trained scientist."[10] In our field this basic point applies with particular force. In making these judgments about the meaning of evidence, one brings a whole sensibility to bear—a sense one develops over the years about how the

political system works, about how honest political leaders are when talking about basic issues in private, about what rings true and what is being said for tactical purposes.

So I doubt very much the second and third guidelines will be of any great value, and I can easily imagine them being counter-productive. This does not mean, of course, that we don't have to worry about the problems that the authors of these guidelines were concerned with. How then should those problems be dealt with? How, in particular, should the problem of cherry-picking be dealt with?

To begin with, we should consider how the current system works. We maintain standards basically by encouraging scholarly debate. People often criticize each others' arguments; those debates help determine the prevailing view. It is not as though we are like trial lawyers, using every trick in the book to win an argument. We use a more restrained version of the adversarial process, where a common interest in getting things right can actually play a major role in shaping outcomes.

This is all pretty standard, but there is one area here where I think a change in norms would be appropriate. This has to do with the prevailing bias against purely negative arguments, and with the prevailing assumption that an author should test his or her own theories.

It is very common in the political science literature to see an author lay out his or her own theory, present one or two alternatives to it, and then look at one or more historical cases to see which approach holds up best. And, lo and behold, the author's own theory always seems to come out on top. We're all supposed to pretend that the author's obvious interest in reaching that conclusion did not color the analysis in any way. But that pose of objectivity is bound to be somewhat forced: true objectivity is simply not possible in such a case. I personally would prefer it if the author just presented the theory, making as strong a case for it as possible, and did not pretend that he or she was "testing" it against its competitors. I would then leave it to others to do the "testing"—and that means that others should be allowed to produce purely negative arguments. If the "test" shows that the theory does not stand up, the analyst should be allowed to stop there. He or she should not be told (as critics often are) that a substitute theory needs to be produced. So if I were coming up with a list of rules for journal editors, I would be sure to include this one.

References

American Political Science Association, Committee on Ethics, Rights and Freedoms. 2012. *A Guide to Professional Ethics in Political Science*, 2nd edition. Washington: APSA. (http://www.apsanet.org/Portals/54/APSA%20Files/publications/ethicsguideweb.pdf, last accessed 7/1/2015).

Gartzke, Erik and Yonatan Lupu. 2012. "Still Looking for Audience Costs." *Security Studies* vol. 21, no. 3: 391–397. (http://dx.doi.org/10.1080/09636412.2012.706486, last accessed 6/27/2015)

Kuhn, Thomas. 1970. "Reflections on My Critics." In Imre Lakatos and Alan Musgrave, eds., *Criticism and the Growth of Knowledge*. Cambridge: Cambridge University Press, 231–278.

Moravcsik, Andrew. 2014. "Transparency: The Revolution in Qualitative Research." *PS: Political Science and Politics* vol. 47, no. 1:

---

[8] See Trachtenberg 2006, ch.4 (https://www.sscnet.ucla.edu/polisci/faculty/trachtenberg/cv/chap4.pdf).

[9] Gartzke and Lupu 2012, 393 n7, 394f.

[10] Kuhn 1970, 233. Note also Stephen Toulmin's reference to the "judgment of authoritative and experienced individuals" (1972, 243).

48–53. (http://dx.doi.org/10.1017/S1049096513001789, last accessed 6/27/2015)

Trachtenberg, Marc. 2005. "The Iraq Crisis and the Future of the Western Alliance," in *The Atlantic Alliance Under Stress*, edited by David M. Andrews. Cambridge: Cambridge University Press. Extended version with links to sources at http://www.sscnet.ucla.edu/polisci/faculty/trachtenberg/useur/iraqcrisis.html.

———. 2006. *The Craft of International History: A Guide to Method*. Princeton University Press. (http://graduateinstitute.ch/files/live/sites/iheid/files/sites/international_history_politics/users/stratto9/public/Trachtenberg%20M%20The_Craft_of_International_His.pdf)

———. 2008. "The United States and Eastern Europe in 1945: A Reassessment." *Journal of Cold War Studies* vol. 10, no. 4: 94–132. (http://dx.doi.org/10.1162/jcws.2008.10.4.94, last accessed 6/27/2015)

———. 2011. "The French Factor in U.S. Foreign Policy during the Nixon-Pompidou period, 1969–1974." *Journal of Cold War Studies* vol. 13, no. 1: 4–59. Extended version with links to sources at https://www.sscnet.ucla.edu/polisci/faculty/trachtenberg/ffus/FrenchFactor.pdf.

———. 2012. "Audience Costs: An Historical Analysis." *Security Studies* vol. 21, no. 1: 3–42. Long version at http://www.sscnet.ucla.edu/polisci/faculty/trachtenberg/cv/audcosts(long).doc.

———. 2013a. "Dan Reiter and America's Road to War in 1941." *H-Diplo/ISSF Roundtable* vol. 5, no. 4 (May). Extended version with links to sources at http://www.polisci.ucla.edu/faculty/trachtenberg/cv/1941.doc.

———. 2013b. "Audience Costs in 1954?" *H-Diplo/ISSF*, September. Extended version with links to sources at http://www.sscnet.ucla.edu/polisci/faculty/trachtenberg/cv/1954(13)(online).doc.

———. 2013c. "Kennedy, Vietnam and Audience Costs." Unpublished manuscript, November. Extended version with links to sources at http://www.sscnet.ucla.edu/polisci/faculty/trachtenberg/cv/viet(final).doc.

Toulmin, Stephen. 1972. *Human Understanding*. vol.1. Princeton: Princeton University Press.

Transparency In Field Research

Transparent Explanations, Yes. Public Transcripts and Fieldnotes, No: Ethnographic Research on Public Opinion

Katherine Cramer
University of Wisconsin, Madison

I am a scholar of public opinion. My main interest is in examining how people understand, or interpret, politics. For that reason, much of my work involves listening to people talk with others with whom they normally spend time. When I listen to the way they sort out issues together, I am able to observe what they perceive to be important, the narratives they use to understand their world, the identities that are central to these understandings, and other important ingredients of public opinion.

My work is therefore primarily qualitative, and usually interpretivist. By interpretivist, I mean that I am trying to capture how people perceive or attribute meaning to their worlds. I treat that effort as a necessary part of trying to understand why they express the opinions that they do.

Across the course of my career, transparency has been a professional necessity. My methods are rather unusual in the field of public opinion research, so the burden is on me to teach my readers and my reviewers what I am doing, why I am doing it, and how my work should be judged. Usually, public opinion scholars focus on individuals' preferences and how to predict them, not on the process of understanding. In addition, we tend to be well versed in the strengths and weaknesses of polling data, but basically unfamiliar with conversational data.

Put another way, reviewers are likely to dive into my papers looking for the dependent variable and the strength of evidence that my results can be generalized to a national population. But my papers usually do not provide information on either of these things. Unless I explain why my work has different qualities to be judged, the typical reviewer will quickly tune out and give the paper a resounding reject after the first few pages.

So the first thing I have had to be transparent about is the fact that much of my work is not attempting to predict preferences. My work typically does not describe how a set of variables co-vary with one another to bring about particular values on a dependent variable. Indeed, I'm not usually talking about causality. These characteristics are just not what scholars typically come across in political science public opinion research. I have had to go out of my way to explain that my work uses particular cases to help explain in detail the process of a group of people making sense of politics. I have had to be up front about the fact that my goal is to help us understand how it is that certain preferences are made obvious and appropriate when objective indicators about a person's life would suggest otherwise.

For example, in a piece I published in the *APSR* in 2012,[1] reviewers helpfully pointed out that I had to bluntly state that my study used particular cases to study a broader question. In short, that article reported the results of a study in which I invited myself into conversations among groups of people meeting in gathering places like gas stations and cafés in communities throughout Wisconsin, especially small towns and rural places, so that I could better understand how group consciousness might lead people to support limited government

Katherine Cramer is Professor of Political Science and Director of the Morgridge Center for Public Service at the University of Wisconsin-Madison. She is online at kathy.cramer@wisc.edu and https://faculty.polisci.wisc.edu/kwalsh2/.

[1] Cramer Walsh 2012.

when their objective interests might suggest otherwise. I had to take several paragraphs to contrast what I was doing against the more familiar positivist approaches. I wrote:[2]

> My purpose in investigating what people say in the groups they normally inhabit in a particular set of communities within one state is to better explain how the perspectives people use to interpret the world lead them to see certain stances as natural and right for people like themselves (Soss 2006, 316). It is motivated by the interpretivist goal of providing a "coherent account of [individuals'] under-standings as a prerequisite for adequate explanation" (Soss 2006, 319; see also Adcock 2003). In other words, to explain why people express the opinions that they do, we need to examine and describe how they perceive the world. In this article I explain the contours of the rural consciousness I observed and then specify its particularity by contrasting it with conversations among urban and suburban groups. That is, this is a constitutive analysis (an examination of what this thing, rural consciousness, consists of and how it works) versus a causal analysis (e.g., an examination of whether living in a rural place predicts rural consciousness—McCann 1996; Taylor 1971; Wendt 1998). The point is not to argue that we see consciousness in rural areas but not in other places, nor to estimate how often it appears among rural residents, nor to describe what a population of people thinks. Instead, the purpose here is to examine what this particular rural consciousness is and what it does: how it helps to organize and integrate considerations of the distribution of resources, decision-making authority, and values into a coherent narrative that people use to make sense of the world. This is not a study of Wisconsin; it is a study of political understand ing and group consciousness that is conducted in Wisconsin (Geertz 1973, 22).

To clarify the stakes, contributions, and implications of this study, allow me to contrast it with positivist approaches. I examine here how people weave together place and class identities and their orientations to government and how they use the resulting perspectives to think about politics. A positivist study of this topic might measure identities and orientations to government, and then include them as independent variables in a multivariate analysis in which the dependent variable is a policy or candidate preference. Such an approach is problematic in this case in the following ways. The positivist model specification assumes that values on one independent variable move independent of the other. Or if using an interaction term, it assumes that people with particular combinations of these terms exhibit a significantly different level of the dependent variable. However, the object of study, or my dependent variable in positivist terms, is not the position on an attitude scale. It is instead the perspectives that

people use to arrive at that position. My object is not to understand the independent effects of identities and attitudes such as trust, or how people with different combinations of these compare to others, but to understand how people themselves combine them—how they constitute perceptions of themselves and use these to make sense of politics.

I include this excerpt to underscore that transparency in the sense of explaining in detail my data collection and analysis procedures, as well as my epistemological approach, has been a professional necessity for me. Without providing extensive detail on my research approach, many readers and reviewers would not recognize the value of my work. Indeed, on one occasion in which I did not take the space to provide such detail, one exceptionally uncharitable reviewer wrote, "I found this chapter very disappointing. This perhaps reflects my bias as a researcher who does formal models and quantitative analysis of data. Briefly, I, too, can talk to taxi drivers or my mother's bridge crowd." There were just 3 additional sentences to this person's review.

Transparency has also been important for me for another reason. In interpretive work, we make evidence more valuable and useful the more context we provide. As we describe and examine the meaning that people are making out of their lives, we better equip our readers to understand and judge our claims the more information we provide about what leads to our interpretations. Positivist work has the burden of providing evidence that a given sample is representative of a broader population. Interpretivists must provide enough context that our interpretations are "embedded in, rather than abstracted from, the settings of the actors studied."[3] Transparency is an integral part of the presentation of our results. For example, in a forthcoming book based on the study I describe above,[4] I explain the physical nature of the settings in which I met with rural residents, the lush green fields and blue skies of the communities in which they resided, what they wore, and how they responded to me as a smiley city girl arriving to their gas stations, etc., in my Volkswagen Jetta to convey in part the complexity of the perspectives with which the people I spent time with viewed the world. They were deeply proud of their communities and their groups of regulars, and at the same time resentful of the economic situations in which they found themselves.

In those respects, I value transparency. But in another respect, I do not. In particular, I do not consider making my field notes publicly available to be a professional duty or necessity. I am troubled by the recent push in our discipline to make available the transcripts of the conversations I observe and my fieldnotes about them to anyone and what it means for my future research. I have three specific concerns.

First, asking me to make my data publicly available assumes that any scholar could use it in the form it exists on my computer. That is, the assumption is that if I provide the tran-

---

[2] Cramer Walsh 2012, 518.

[3] Schwartz-Shea and Yanow 2012, 47.
[4] Cramer 2016.

scripts of the conversations, and a key providing an explanation of important characteristics of each speaker, any scholar would be able to treat those words as representations of the conversation. I am just not confident that is possible. My fieldnotes are pretty detailed, but I do not note the tone or inflection of every sentence. I do not include every detail of the town or the venue in which I visit with people. I record details about what people are wearing, how the coffee tasted, and the pitch of some folks' voices, but when I re-read a transcript, a thousand images, sounds, and smells enter my mind to round out the impression of the messages people were conveying to me and to each other. For this reason, I do not send assistants to gather data for me. I need to be there. I need to experience the conversation in all its fullness to get the best possible picture of where people are coming from.

Could a scholar still get something useful from my transcripts? Perhaps—the frequency of the use of some words, the general frames in which certain issues are talked about—all of that would be available in my transcripts. But my sense is that this push for transparency is not about broadening the pool of available data for people. It is about the assumption that making data publicly available will enable replication and the ability to more easily assess the validity of an argument. It is driven by the assumptions that the best social science is social science that is replicable and consists of conclusions that another scholar, given the same data, would arrive at independently. But I do not agree that the transcripts that I or my transcribers have typed out, and the accompanying memos that I have written would allow replication, nor would I expect another scholar to necessarily reach the same conclusions based on re-reading my transcripts.

Let me put it another way. The field of political behavior is pretty well convinced that Americans are idiots when it comes to politics. When you read the transcripts of the conversations I observe, it is not difficult to come away with the conclusion that those exchanges support that general conclusion. The words people have said, typed out on a page, do seem ignorant according to conventional standards. However, when the conversational data is combined with the many intangible things of an interaction—the manner in which people treat each other, the physical conditions in which they live—their words take on meaning and reasonableness that is not evident from the transcription of their words alone.

Let me reiterate that the purpose of my work is not to estimate coefficients. There is not a singular way to summarize the manner in which variables relate to one another in my data. Instead, what I do is try to characterize in as rich a manner possible how people are creating contexts of meaning together. Should I enable other scholars to look at my data to see if they would reach the same conclusion? I do not think it is possible to remove me from the analysis. In my most recent project, conversations about the University of Wisconsin-Madison became a key way for me to examine attitudes about government and public education. I work at that institution. I have a thick Wisconsin accent. My presence became a part of the data. Another scholar would not have all of the relevant data

needed for replication unless he or she is me.

I am against this push for making transcripts and fieldnotes publicly available for another reason. Ethically, I would have a very difficult time inviting myself into conversations with people if I knew that not only would *I* be poring over their words in detail time and time again, but that an indeterminate number of other scholars would be doing so as well, in perpetuity. How do we justify that kind of invasion? You might say that survey research does this all the time—survey respondents are giving the permission for an indeterminate number of people to analyze their opinions for many years to come. But we tell respondents their views will be analyzed in the aggregate. Also, collecting conversational data, in person, in the spaces that people normally inhabit, with the people they choose of their own volition to do so, is not the same as collecting responses in a public opinion poll, even if that poll is conducted in person. When you are an interviewer for a public opinion poll, you are trained to be civil, but basically nondescript—as interchangeable with other interviewers as possible. That is just about the opposite of the approach needed in the kind of research I conduct. I have to walk into a group as my authentic self when I ask if I can join their coffee klatch and take up some other regular's barstool or chair. I am able to do the kind of work I do because I am willing to put myself out there and connect with people on a human level. And I am able to gather the data that I do because I can tell people verbally and through my behavior that what they see is what they get. If the people I studied knew that in fact they were not just connecting with me, but with thousands of anonymous others, I would feel like a phony, and frankly would not be able to justify doing this work.

My main worry with this push for transparency is that it is shooting our profession in the foot. I am concerned in particular about the push for replicability. Does the future of political science rest on replicability? I have a hard time seeing that. Perhaps because I work at the University of Wisconsin-Madison, in which tenure and public higher education are currently under fire, when I think about the future of our profession, I think about the future of higher education in general. It seems to me that our profession is more likely to persist if we more consciously consider our relevance to the public and how our institutions might better connect with the public. I am not saying that we should relax scientific standards in order to appease public will. I am saying that we should recognize as a discipline that there are multiple ways of understanding political phenomena and that some ways of doing so put us in direct connection with the public and would be endangered by demanding that we post our transcripts online. Why should we endanger forms of data collection that put us in the role of ambassadors of the universities and colleges at which we work, that put us in the role of listening to human beings beyond our campus boundaries? It is not impossible to do this work while making the transcripts and field notes publicly available, but it makes it much less likely that any of us will pursue it.

I do not think outlawing fieldwork or ethnography is the point of the data access initiative, but I fear it would be an

unintended consequence resulting from a lack of understanding of interpretive work—the kind of misunderstanding that leads some political scientists to believe that all we are up to is talking with taxi drivers.

My interpretive work seeks to complement and be in dialogue with positivist studies of public opinion and political behavior. Its purpose is to illuminate the meaning people give to their worlds so that we can better understand the political preferences and actions that result. Understanding public opinion requires listening to the public. Transparency in this work is essential to make my methods clear to other scholars in my field who typically are unfamiliar with this approach, so that they can understand and judge my arguments. But I do not think that mandated transparency should extend to providing my transcripts and fieldnotes. My transcripts and fieldnotes are not raw data. The raw data exist in the act of spending time with and listening to people. That cannot be archived. The expectation for interpretive work should be that scholars thoroughly communicate their methods of data collection and analysis and provide rich contextual detail, including substantial quoting of the dialogue observed. There are many excellent models of such transparency in interpretive ethnographic work already in political science, which we can all aspire to replicate.[5]

References

Adcock, Robert. 2003. "What Might It Mean to Be an 'Interpretivist'?" *Qualitative Methods: Newsletter of the American Political Science Association Organized Section* vol. 1, no. 2: 16–18.

Cramer, Katherine. 2016. *The Politics of Resentment: Rural Consciousness in Wisconsin and the Rise of Scott Walker*. Chicago: University of Chicago Press.

Cramer Walsh, Katherine. 2012. "Putting Inequality in Its Place: Rural Consciousness and the Power of Perspective." *American Political Science Review* vol. 106, no. 3: 517–532.

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

McCann, Michael. 1996. "Causal versus Constitutive Explanations (or, On the Difficulty of Being So Positive . . .)." *Law and Social Inquiry* vol. 21, no. 2: 457–482.

Pachirat, Timothy. 2011. *Every Twelve Seconds: Industrialized Slaughter and the Politics of Sight*. New Haven: Yale University Press.

Schwartz-Shea, Peregrine, and Dvora Yanow. 2012. *Interpretive Research Design: Concepts and Processes*. New York: Routledge.

Simmons, Erica. 2016. *The Content of their Claims: Market Reforms, Subsistence Threats, and Protest in Latin America*. New York: Cambridge University Press, forthcoming.

Soss, Joe. 2000. *Unwanted Claims: The Politics of Participation in the U.S. Welfare System*. Ann Arbor: University of Michigan Press.

Soss, Joe. 2006. "Talking Our Way to Meaningful Explanations: A Practice-Centered Approach to In-depth Interviews for Interpretive Research." In *Interpretation and Method*, edited by Dvora Yanow and Peregrine Schwartz-Shea. New York: M. E. Sharpe, 127–149.

---

[5] See, e.g., Pachirat 2011; Simmons 2016; Soss 2000.

Taylor, Charles. 1971. "Interpretation and the Sciences of Man." *Review of Metaphysics* vol. 25, no. 1: 3–51.

Wendt, Alexander. 1998. "On Constitution and Causation in International Relations." *Review of International Studies* vol. 24, no. 5: 101–118.

Research in Authoritarian Regimes: Transparency Tradeoffs and Solutions

Victor Shih
University of California, San Diego

Conducting research in authoritarian regimes, especially ones with politicized courts, bureaucracy, and academia, entails many risks not encountered in research in advanced democracies. These risks affect numerous aspects of research, both qualitative and quantitative, with important implications for research transparency. In this brief essay, I focus on the key risk of conducting research in established authoritarian regimes: namely, physical risks to one's local informants and collaborators. Minimizing these risks will entail trading off ideal practices of transparency and replicability. However, scholars of authoritarian regimes can and should provide information on how they have tailored their research due to constraints imposed by the regime and their inability to provide complete information about interviewees and informants. Such transparency at least would allow readers to make better judgments about the quality of the data, if not to replicate the research. Also, scholars of authoritarian regimes can increasingly make use of nonhuman data sources that allow for a higher degree of transparency. Thus, a multi-method approach, employing data from multiple sources, is especially advisable for researching authoritarian regimes.

First and foremost, conducting research *in* authoritarian countries can entail considerable physical risks to one's research subjects and collaborators who reside in those countries. To the extent that authorities impose punitive measures on a research project, they are often inflicted on in-country research subjects and collaborators because citizens in authoritarian countries often do not have legal recourse. Thus a regime's costs of punishing its own citizens are on average low relative to punishing a foreigner. At the same time, the deterrence effect can be just as potent. Thus, above all else, researchers must protect subjects and collaborators as much as possible when conducting research in authoritarian regimes, often to the detriment of other research objectives.

For example, academics who conduct surveys in China often exclude politically sensitive questions in order to protect collaborators. The local collaborators, for their part, are careful and often have some idea of where the "red line" of unacceptable topics is. Beyond the judgment of the local collaborators,

---

Victor Shih is associate professor at the School of Global Policy and Strategy at the University of California, San Diego. He is online at vcshih@ucsd.edu, http://gps.ucsd.edu/faculty-directory/victor-shih.html and on twitter @vshih2.

however, a foreign academic additionally should pay attention to the latest policy trends. When necessary, the foreign academic should use an additional degree of caution, especially when the authorities have signaled heightened vigilance against "hostile Western forces," as China has done recently. This degree of caution will necessarily exclude certain topics from one's research agenda, which is detrimental to academic integrity. It is an open secret in the China studies community that China's authoritarian censors have a deep impact on the choice of research topics by China scholars located in the United States and other democracies. Researchers who work on certain "red line" topics will not and *should not* collaborate with mainland Chinese academics, and often time, they are barred from entering China.[1] Even when one is working on an "acceptable" topic, instead of asking sensitive political questions in a survey, for example, academics must ask proxy questions that are highly correlated with the sensitive questions. This is often done at the advice of the local collaborators, who are indispensable for conducting survey research in China. Prioritizing the safety of one's collaborator will inevitably sacrifice some aspects of the research such as excluding certain topics and avoiding certain wordings in a questionnaire.

In current scholarly practice, such tactical avoidance of topics and wording in one's research is kept implicit in the final product. At the most, one would mention such limitations during a talk when someone asks a direct question. It would be a welcome step toward greater research transparency if researchers who have had to, for example, change the wording of questions in a survey due to fear of censure would outline in footnotes what the ideal wordings may have been, as well as the reason for changing the wording. We now have access to several leaked documents from the Chinese Communist Party specifying exactly the type of "hostile" Western arguments that are being watched carefully by the regime.[2] Such official documents provide a perfectly legitimate reason for researchers not to trespass certain "red lines" topic in their works. Making explicit the distortions caused by fear of punishment from the regime also helps the readers adjudicate the validity and limitations of the results. And such disclosures may even motivate other scholars to find ways to overcome these distortions with new methods and data.

Because political processes in authoritarian regimes often remain hidden, interviews continue to be a key step in researching these regimes. Interviews are especially important in the hypothesis-generating stage because the incentives facing political actors in various positions in the regime are often quite complex. Thus, extended conversations with informants are the best way to obtain some clarity about the world in which they live. However, foreign academics must take an even more cautious approach with interviews than with large-sample surveys because even grassroots informants can get in trouble

for revealing certain information to foreign academics. In China, for example, local officials are evaluated by their ability to "maintain stability."[3] Thus, foreign researchers conducting research on collective action may deeply alarm the local authorities. At the same time, arresting or detaining a foreign researcher may draw Beijing's attention to local problems, which would be detrimental to local officials. Therefore, local authorities' best response is to mete out punishment on local informants. Such a scenario must be avoided as much as possible.

In conducting interview research in authoritarian regimes, the setting in which one conducts interviews may greatly affect the quality of the interview. An introduction through mutual friends is the best way to meet an informant. Randomly talking to people on the street, while necessary at times, may quickly draw the attention of the authorities. Again, using China as an example, nearly all urban residential communities have guards and residential committee members who watch over the flow of visitors.[4] Similarly, guards and informants are common in one's workplace. Instead of barging into the informant's abode or workplace, which can draw the attention of residential or workplace monitors, a mutual friend can arrange for a casual coffee or dinner somewhere else. The informant will often be at ease in this setting, and there is less of a chance for the residential monitors to sound the alarm. An introduction by friends is equally useful for elite interviews for similar reasons. Because the circumstances of the interviews can greatly affect the subject's ease of conversation and the quality of the information, researchers in an authoritarian regime may want to provide readers with information on interview settings. At the same time, they may not be able to live up to standards of disclosure that are taken for granted for research conducted in Western democracies: The more a researcher relies upon mutual friends and other personal connections, the greater the caution a researcher needs to take not to disclose information that would put not only the informant at risk but also the person who put the researcher in touch with the informant.

As a consequence, when academics are citing interview notes in their writing, they must obfuscate or even exclude key details about the informants such as their geographic location, the date of the interviews, and specific positions in an organization. Under most circumstances, I would even advise against sharing field notes with other researchers or depositing them in a data archive.

This said, one should take meticulous notes when interviewing informants in authoritarian countries. In the interest of research transparency, researchers should provide extensive information and discuss to what extent informants in various positions are able to provide unique insights that help the hypothesis generation or testing processes. As discussed below, the reporting of the informants' positions should leave out potentially identifying information, but should nonetheless be sufficiently informative to allow readers to judge whether such a person may be in a position to provide credible information to the researcher. For example, one can report "planning

---

[1] de Vise 2011.

[2] See, for example, Central Committee of the Chinese Communist Party. 2014. Communiqué on the Current State of the Ideological Sphere (Document 9). Posted on http://www.chinafile.com/document-9-chinafile-translation (last accessed 7/4/2015).

[3] Edin 2003, 40.

[4] Read 2012, 31–68.

official in X county, which is near a major urban center" or "academic with economics specialization at a top 50 university"…etc. Although this level of disclosure is far from perfect, it at least provides readers with some sense about the qualification of the informant in a given topic area.

Special care to protect subjects' identities also must be taken during the research process and in the final products. For one, to the extent that the researcher is keeping a list of informants, that list should be stored separately from the interview notes. Both the list of informants and the interview notes should be stored in a password-protected setting, perhaps in two secure clouds outside of the authoritarian country. At the extreme, the researcher may—while in the country—need to keep only a mental list of informants and only noting the rough positions of the informants in written notes. In China, for example, foreign academics have reported incidents of the Chinese authorities breaking into their hotel rooms to install viruses and spyware on their computers. Thus, having a password-protected laptop is far from sufficient. When academics are citing interview notes in their writing, they must obfuscate or even exclude key details about the informants such as their geographic location, the date of the interviews, and specific positions in an organization.

To be sure, these rules of thumb go against the spirit of data access and research transparency. They also make pure replication of qualitative interview data-collection impossible. At most, other scholars may be able to interview informants in similar positions but likely in different geographical locations, and subsequent interviews may yield totally different conclusions. However, this is a tradeoff that researchers of authoritarian regimes must accept without any leeway. Because informants in authoritarian regimes can face a wide range of physical and financial harm, their safety must come first before other research criteria.

Although researchers of authoritarian regimes cannot provide complete transparency in their interview data, they can compensate with a multi-method, multi-data approach that provides a high degree of transparency for other non-human sources of data. Increasingly, researchers who glean some key insights from interviews are also testing the same hypotheses using nonhuman quantitative data such as remote-sensing data,[5] economic and financial data,[6] textual data,[7] and elite biographical data.[8] These datasets are typically collected from publicly available sources such as the Internet or satellite imageries and made widely available to other researchers for replication purposes.[9] Instead of only relying on somewhat secretive interview data, researchers of authoritarian regimes can increasingly make full use of other data sources to show the robustness of their inferences. This does not mean that interviews are no longer needed because nothing can quite replace interviews in the initial hypothesis generation stage. It does

mean that hypothesis-testing has become much less "black boxy" for empirical research of authoritarian regimes.

References

de Vise, Daniel. 2011. "U.S. Scholars Say Their Book on China Led to Travel Ban." *Washington Post*, 20 August.

Edin, Maria. 2003. "State Capacity and Local Agent Control in China: CCP Cadre Management from a Township Perspective." *China Quarterly* vol. 173 (March):35–52.

King, Gary, Jennifer Pan, and Margaret Roberts. 2013. "How Censorship in China Allows Government Criticism but Silences Collective Expression." *American Political Science Review* vol. 107, no. 1:1–18.

Mattingly, Daniel. 2015. "Informal Institutions and Property Rights in China." Unpublished manuscript, University of California, Berkeley.

Read, Benjamin Lelan. 2012. *Roots of the State: Neighborhood Organization and Social Networks in Beijing and Taipei*. Stanford: Stanford University Press.

Shih, Victor, Christopher Adolph, and Liu Mingxing. 2012. "Getting Ahead in the Communist Party: Explaining the Advancement of Central Committee Members in China." *American Political Science Review* vol. 106, no. 1: 166–187.

Shih, Victor, Wei Shan, and Mingxing Liu. 2008. "Biographical Data of Central Committee Members: First to Sixteenth Party Congress." Retrievable from http://faculty.washington.edu/cadolph/?page=61 (last accessed 7/5/2015).

Wallace, Jeremy L. 2015. "Juking the Stats? Authoritarian Information Problems in China." *British Journal of Political Science*, forthcoming; available at http://dx.doi.org/10.1017/S0007123414000106.

Transparency in Intensive Research on Violence: Ethical Dilemmas and Unforeseen Consequences

Sarah Elizabeth Parkinson
University of Minnesota

Elisabeth Jean Wood
Yale University

Scholars who engage in intensive fieldwork have an obligation to protect research subjects and communities from repercussions stemming from that research.[1] Acting on that duty

---

[1] We generally take "intensive" fieldwork to mean fieldwork that is qualitative and carried out during long-term (six months or more), at least partially immersive stays in the field, incorporating methods such as participant observation, in-depth interviewing, focus groups,

---

[5] Mattingly 2015.

[6] Wallace 2015.

[7] King et al. 2013.

[8] Shih et al. 2012.

[9] See, e.g., Shih et al. 2008.

not only paves the way towards ethical research but also, as we argue below, facilitates deeper understanding of people's lived experiences of politics. For scholars who study topics such as violence, mobilization, or illicit behavior, developing and maintaining their subjects' trust constitutes the ethical and methodological foundation of their ability to generate scholarly insight. Without these commitments, work on these topics would not only be impossible; it would be unethical.

As scholars with extensive fieldwork experience, we agree with the principle of research transparency—as others have noted,[2] no one is *against* transparency. However, we find current efforts in the discipline to define and broadly institutionalize particular practices of transparency and data access, embodied in the DA-RT statement,[3] both too narrow in their understanding of "transparency" and too broad in their prescriptions about data access.

In this essay, we advance four arguments. First, there are meanings of "transparency" at the core of many field-based approaches that the initiative does not consider.[4] Second, standards governing access to other kinds of data should not in practice and could not in principle apply to projects based on intensive fieldwork: "should not in practice" in order to protect human subjects; "could not in principle" because of the nature of the material gathered in such research. Third, while we support the aim of researcher accountability, a frequently advocated approach—replicability—while central to some research methods, is inappropriate for scholarship based on intensive fieldwork, where accountability rests on other principles. Fourth, the implementation of a disciplinary norm of data access would undermine ethical research practices, endanger research participants, and discourage research on important but challenging topics such as violence.

We illustrate the issues from the perspective of research on or in the context of violence (hereafter "violence research"). Our emphasis on ethics, our views on empirical evidence and its public availability, and our concerns regarding emergent conflicts of interest and problematic incentive structures are relevant to scholars working in an array of sub-fields and topics, from race to healthcare.

Transparency in Research Production and Analysis

We agree with the general principles of production and analytic transparency: authors should clearly convey the research procedures that generate evidence and the analytical processes that produce arguments. Those conducting violence research necessarily situate these principles within broader discussions of trust, confidentiality, and ethics. When field researchers think about transparency, they think first of their relationships

with, disclosures to, and obligations towards participants and their communities.[5]

The values of beneficence, integrity, justice, and respect that form the cornerstones of what is broadly referred to as "human subjects research"[6] are put into practice partially, though not exclusively, via the principles of informed consent and "do no harm." Informed consent is fundamentally a form of transparency, one that DA-RT does not address. In its simplest form, informed consent involves discussing the goals, procedures, risks, and benefits of research with potential participants. Because the possible effects of human subjects research include what institutional review boards (IRBs) rather clinically term "adverse events" such as (re)traumatization, unwanted public exposure, and retaliation, responsible researchers spend a considerable amount of time contemplating how to protect their subjects and themselves from physical and psychological harm. Most take precautions such as not recording interviews, encrypting field notes, using pseudonyms for both participants and field sites, embargoing research findings, and designing secure procedures to back up their data. In the kind of research settings discussed here, where research subjects may literally face torture, rape, or death, such concerns must be the first commitment of transparency, undergirding and conditioning all other considerations.[7]

Transparency is closely related to trust. Those conducting intensive fieldwork understand trust as constructed through interaction, practice, and mutual (re)evaluation over time. Trust is not a binary state (e.g., "no trust" versus "complete trust") but a complex, contingent, and evolving relationship. Part of building trust often involves ongoing discussions of risk mitigation with research subjects. For example, during field work for a project on militant organizations,[8] Parkinson's Palestinian interlocutors in Lebanon taught her to remove her battery from her mobile phone when conducting certain interviews, to maintain an unregistered number, and to buy her handsets using cash. They widely understood mobile phones to be potential listening and tracking devices.[9] The physical demonstration of removing a mobile battery in front of her interlocutors showed that she understood the degree of vulnerability her participants felt, respected their concerns, and would not seek to covertly record interviews. Over time, as Parkinson's interlocutors observed her research practices through repeated interactions, experienced no adverse events, read her work, and felt that their confidentiality had been respected, they became increasingly willing to share more sensitive knowledge.

We and other scholars of violence have found that participants come to trust the researcher not just to protect their identities, but also to use her judgment to protect them as unforeseen contingencies arise. While having one's name or organization visible in one context may provide some measure

---

and community mapping. We use the terms "subjects," "participants," and "interlocutors" interchangeably.

[2] Pachirat 2015; Isaac 2015.

[3] DA-RT 2014.

[4] Isaac (2015, 276) asks "whether the lack of transparency is really the problem it is being made out to be," by DA-RT, a concern we share.

[5] Wood 2006; Thomson 2010.

[6] Office of the Secretary of Health and Human Services 1979.

[7] For a more in-depth discussion, please see Fujii 2012.

[8] Parkinson 2013a.

[9] Parkinson 2013b, Appendix C.

of protection or status, in others it may present significant risk. And "context" here may change rapidly. For example, scholars working on the Arab Uprisings have noted that activists who initially and proudly gave researchers permission to quote them by name were later hesitant to speak with the same scholars due to regime changes and shifts in overall political environment.[10] There is often no way to know whether an activist who judges herself to be safe one day will be criminalized tomorrow, next month, or in five years. Those in this position may not be able to telephone or email a researcher in order to remove their name from a book or online database; they may not know until it is too late.

In the more general realm of "production transparency," field-intensive research traditions broadly parallel many other methodologies. The best work explains why and how field sites, populations, interview methods, etc. fit within the research design in order for the reader to evaluate its arguments. Many of these field-based methods (e.g., participant observation) also require researchers to evaluate how elements of their background and status in the field affect their interactions with participants and their analysis. Reflexivity and positionality, as these techniques are termed, thus fundamentally constitute forms of transparency.[11]

Thus we suggest that the principle of production transparency should be informed both by human subject concerns—particularly in the case of violence research—and by the nature of the evidence that intensive fieldwork generates.

Turning to analytic transparency, we agree: an author should convey the criteria and procedures whereby she constructed her argument from the evidence gathered. For research based on extensive fieldwork this might mean, for example, being explicit about why she weighed some narratives more heavily than others in light of the level of detail corroborated by other sources, the length of the researcher's relationship to the participant, or the role of meta-data.[12] Furthermore, the author should be clear about the relationship between the field research and the explanation: did the scholar go to the field to evaluate alternative well-developed hypotheses? To construct a new theory over the course of the research? Or did the research design allow for both, with later research evaluating explicit hypotheses that emerged from a theory drawing on initial data?

Data Access

The values of beneficence, integrity, justice, and respect imply not only that participants give informed consent but also that their privacy be protected. For some types of research, maintaining subject confidentiality may be easily addressed by posting only fully de-identified datasets. But in the case of intensive field research, "data" can often not be made available for methodological reasons, and in the case of violence research, it should almost always *not* be made accessible for ethical reasons.

The very nature of such empirical evidence challenges the principle of data access. Evidence generated through participant observation, in-depth interviews, community mapping, and focus groups is deeply relational, that is, constructed through the research process by the scholar and her interlocutors in a specific context. As other authors in this newsletter underscore,[13] these materials do not constitute "raw data." Rather, they are recordings of intersubjective experiences that have been interpreted by the researcher.

We add that the idea of writing field notes or conducting interviews with the anticipation of making sensitive materials available would fundamentally change the nature of interaction with subjects and therefore data collection. Among other problems, it would introduce components of self-censorship that would be counterproductive to the generation of detailed and complete representations of interactions and events. Under requirements of public disclosure, violence scholars would have to avoid essential interactions that inform the core of their scholarship. A responsible researcher would not, for example, visit a hidden safe house to conduct an interview with rebel commanders or attend a meeting regarding an opposition party's protest logistics. Any representation of such interactions, if they were to be ethically compiled, would also be unusably thin. More broadly, to imply that all field experiences can and should be recorded in writing and transmitted to others is to deny the importance of *participation* in intensive fieldwork: taking risks, developing trust, gaining consent, making mistakes, sharing lived experiences, and comprehending the privilege of being able to leave.

In some settings, even if researchers remove identifiers—often impossible without rendering materials useless—posting the data would nonetheless do harm to the community and perhaps enable witch hunts. For example, although field site identities are sometimes masked with pseudonyms, they are sometimes very clear to residents and relevant elites. If field notes and interviews are easily accessible, some research participants may fear that others may consequently seek to retaliate against those whom they believe shared information. Whether that belief is correct or not, the damage may be harmful, even lethal, and may "ruin" the site for future research precisely because the so-called "raw" data were made accessible. Posting such data may undermine *perceptions* of confidentiality, and thereby indirectly cause harm.

Nonetheless, on some topics and for some settings, some material can and should be shared. For example, if a scholar records oral histories with subjects who participate with the clear understanding that those interviews will be made public (and with a well-defined understanding about the degree of confidentiality possible, given the setting), the scholar should

---

[10] Parkinson's confidential conversations with Middle East politics scholars, May and June 2015. These conversations are confidential given that several of these researchers are continuing work at their field sites.

[11] See Schatz 2009; Wedeen 2010; Yanow and Schwartz-Shea 2006. See Carpenter 2012; Pachirat 2009; Schwedler 2006 for examples of different approaches.

[12] Fujii 2010.

[13] See, e.g., Cramer 2015; Pachirat 2015.

in general make those materials available. The Holocaust testimonies available through the Yale University Fortunoff Video Archive for Holocaust Testimonies (HVT) and the University of Southern California Shoah Foundation Institute for Visual History provide good examples and have been innovatively employed by political scientists.[14] Even when consent has been granted, however, the scholar should use her own judgment: if posting some sort of transcript might result in harm to the subject, the researcher should consider not making the transcript available, even though she had permission to do so.

The Goals of Research Transparency in Intensive Field Research

As social scientists, field researchers are committed to advancing scholarly understanding of the world. This commitment does not, however, imply that researchers using these approaches thereby endorse a norm of accountability—replicability—appropriate to other methods. What would "replicability" mean in light of the nature of intensive, field-based research?

"Replicability" is often taken to mean "running the same analyses on the same data to get the same result."[15] For some projects on political violence, it is conceivable that this could be done once the data had been codified. For example, presumably a scholar could take the database that Scott Straus built from his interviews with Rwandan *genocidaires* and replicate his analysis.[16] But could she take his transcripts and interview notes and build an identical database? Without time in the field and the experience of conducting the interviews, it is very unlikely that she would make the same analytical decisions. In general, one cannot replicate intensive fieldwork by reading a scholar's interview or field notes because her interpretation of evidence is grounded in her situated interactions with participants and other field experiences.

Without access to data (in fact and in principle), on what grounds do we judge studies based on intensive fieldwork? We cannot fully address the issue here but note that—as is the case with all social science methods—field-intensive approaches such as ethnography are better suited to some types of understanding and inference than others. Scholars in these traditions evaluate research and judge accountability in ways other than replication.[17] The degree of internal validity, the depth of knowledge, the careful analysis of research procedures, the opportunities and limitations presented by the researcher's identity, the scholarly presentation of uncertainties (and perhaps mistakes): all contribute to the judgment of field-intensive work as credible and rigorous.[18] Furthermore,

scholars in these traditions expect that the over-arching findings derived from good fieldwork in similar settings on the same topic should converge significantly. Indeed, scholars are increasingly exploring productive opportunities for comparison and collaboration in ethnographic research.[19]

However, divergence of findings across space or time may be as informative as convergence would be. Failure to "exactly replicate" the findings of another study can productively inform scholarly understanding of politics. Revisits, for example, involve a scholar returning to a prior research site to evaluate the claims of a previous study.[20] The tensions and contradictions that projects such as revisits generate—for example, a female researcher visiting a male researcher's former field site—provide key opportunities for analysis. Divergence in fieldwork outcomes should not necessarily be dismissed as a "problem," but should be evaluated instead as potentially raising important questions to be theorized.

Unforeseen Consequences of DA-RT's Implementation

In addition to the above concerns, we worry that DA-RT's implementation by political science journals may make field research focused on human subjects unworkable. Consider the provision for human subjects research in the first tenet of DA-RT:

> If cited data are restricted (e.g., classified, require confidentiality protections, were obtained under a non-disclosure agreement, or have inherent logistical constraints), authors must notify the editor at the time of submission. The editor shall have full discretion to follow their journal's policy on restricted data, including declining to review the manuscript or granting an exemption with or without conditions. The editor shall inform the author of that decision prior to review.[21]

We are not reassured by the stipulation that it is at the editors' discretion to exempt some scholarship "with or without conditions." There are at least two reasons why it is highly problematic that exemption is granted at the discretion of editors rather than as the rule.

First, confidentiality is an enshrined principle of human subjects research in the social sciences as is evident in the Federal "Common Rule" that governs research on human subjects and relevant documents.[22] To treat confidentiality as necessitating "exemption" thus undermines the foundational principles of human subjects research and would unintentionally constrict important fields of inquiry. The idea that political scientists wishing to publish in DA-RT-compliant journals would either have to incorporate a full public disclosure agreement into their consent procedures (thus potentially hamstring-

---

[14] See, e.g., Finkel 2015.

[15] King 1995, 451 n2. King expressly notes that this process should "probably be called 'duplication' or perhaps 'confirmation'" and that "replication" would actually involve reproducing the initial research procedures.

[16] Straus 2005.

[17] Wedeen 2010; Schatz 2009; Yanow and Schwartz-Shea 2006.

[18] See, e.g., Straus 2005; Wood 2003; Autesserre 2010; Parkinson 2013a; Mampilly 2011; Pachirat 2011; Fujii 2009.

[19] See, e.g. Simmons and Smith 2015.

[20] Burawoy 2003.

[21] DA-RT 2014, 2.

[22] Protection Of Human Subjects, Code of Federal Regulations, TITLE 45, PART 46, Revised January 15, 2009. http://www.hhs.gov/ohrp/policy/ohrpregulations.pdf.

ing the research they wish to conduct) or have to risk later rejection by editors who deny the necessary exemption places fieldworkers in a double bind. Interviewing only those who agree to the resulting transcript's posting to a public archive will restrict the range of topics that can be discussed and the type of people who will participate, thereby fundamentally undermining research on sensitive topics including, but not limited to, political violence.

Moreover, chillingly, this statement places the researcher's incentives and norms at odds with those of her interlocutors. Due to the risk of non-publication, the researcher has a conflict of interest; the reward of publication is available unambiguously only to those who convince interlocutors to agree to digital deposit regardless of the topic of inquiry. But how could the Middle Eastern activists referenced above give *informed* consent for a researcher to publicly deposit interview transcripts, given that they could not possibly know in 2011 that they would be targeted by a future regime in 2015? The answer is that when it comes to some topics, there is really no way to provide such consent given that it is unknown and unknowable what will happen in five years. Moreover, even if they did consent, can the researcher be sure that the interview was fully anonymized? The ethical response for a researcher is to publish research on sensitive topics in journals that do not subscribe to DA-RT principles and in books. However, 25 of the major disciplinary journals have affirmed the DA-RT statement. (One leading APSA journal, *Perspectives on Politics*, has refused to endorse it;[23] others such as *World Politics*, *International Studies Quarterly*, and *Politics and Society* have yet to decide.) The result will be that junior scholars who need publications in major journals will have fewer publication venues—and may abandon topics such as violence for ones that would be easier to place.

These are not abstract concerns. A number of young violence scholars pursuing publication in major disciplinary journals have been asked during the review process to submit confidential human subjects material for verification or replication purposes. Editors and reviewers may be unaware that human subjects ethics and regulations, as well as agreements between researchers and their IRBs, protect the confidentiality of such empirical evidence. To the extent that journal editors and reviewers widely endorse the DA-RT principles, early career scholars face a kind of Sophie's choice between following long-institutionalized best practices designed to protect their subjects (thereby sacrificing their own professional advancement), or compromising those practices in order to get published in leading journals. We trust that journals will come to comprehend that endorsing a narrow understanding of transparency over one informed by the challenges and opportunities of distinct methodological approaches limits the topics that can be ethically published, treats some methods as inadmissible, and forces wrenching professional dilemmas on researchers. But by that time, significant damage may have been done to important lines of research, to academic careers, and to intellectual debate.

[23] Isaac 2015.

Second, the discretion to decide which research projects earn an editor's exemption opens scholars to uninformed decisions by editors and opens authors, reviewers, and editors to moral quandaries. How can we, as a discipline, ask journal editors who come from a broad range of research backgrounds to adjudicate claims regarding the degree of risk and personal danger to research subjects (and to researchers) in a host of diverse local situations? How can a scholar researching a seemingly innocuous social movement guarantee that field notes posted online, won't *later* become the basis for a regime's crackdown? What if a journal editor accepts reviewers' demands that fieldnotes be shared, pressures a junior scholar who needs a publication into posting them, and learns five years down the line that said notes were used to sentence protestors to death? The journal editor's smart choice is not to publish the research in the first place, thus contracting a vibrant field of inquiry.

The ethical default in these situations should be caution and confidentiality rather than "exemption" from mandatory disclosure. The discipline should not construct reward structures that fundamentally contradict confidentiality protections and decontextualize risk assessments.

Conclusion

While DA-RT articulates one vision of transparency in research, it neglects key aspects of transparency and ethics that are crucial to intensive field research and especially to studies of political violence. If applied to intensive field research, blanket transparency prescriptions would undermine the nature of long-established methods of inquiry and institutionalize incentives promoting ethically and methodologically inappropriate research practices. In these settings, DA-RT's requirements may make consent improbable, inadvisable, or impossible; undermine scholarly integrity; and limit the grounded insight often only available via field-intensive methodologies. The stakes are more than academic. In violence research, it is the physical safety, job security, and community status of our research participants that is also at risk.

References

Autesserre, Séverine. 2010. *The Trouble with the Congo: Local Violence and the Failure of International Peacebuilding*. New York: Cambridge University Press.

Burawoy, Michael. 2003. "Revisits: An Outline of a Theory of Reflexive Ethnography." *American Sociological Review* vol. 68, no. 5: 645–679.

Carpenter, Charli. 2012. "'You Talk Of Terrible Things So Matter-of-Factly in This Language of Science': Constructing Human Rights in the Academy." *Perspectives on Politics* vol. 10, no. 2: 363–383.

Cramer, Katherine. 2015. "Transparent Explanations, Yes. Public Transcripts and Fieldnotes, No: Ethnographic Research on Public Opinion" *Qualitative and Multi-Method Research: Newsletter of the American Political Science Association's QMMR Section* vol. 13, no. 1.

"Data Access and Research Transparency (DA-RT): A Joint Statement by Political Science Journal Editors." 2014. At http://media.wix.com/ugd/fa8393_da017d3fed824cf587932534c860ea25.pdf (last accessed 7/10/2015).

Finkel, Evgeny. 2015. "The Phoenix Effect of State Repression: Jewish Resistance During the Holocaust." *American Political Science Review* vol. 109, no. 2: 339–353.

Fujii, Lee Ann. 2009. *Killing Neighbors: Webs of Violence in Rwanda*. Ithaca, NY: Cornell University Press.

———. 2010. "Shades of Truth and Lies: Interpreting Testimonies of War and Violence." *Journal of Peace Research* vol. 47, no. 2: 231–241.

———. 2012. "Research Ethics 101: Dilemmas and Responsibilities." *PS: Political Science & Politics* vol. 45, no. 4: 717–723.

Isaac, Jeffrey C. 2015. "For a More Public Political Science." *Perspectives on Politics* vol. 13, no. 2: 269–283.

King, Gary. 1995. "Replication, Replication." *PS: Political Science & Politics* vol. 28, no. 3: 444–452.

Mampilly, Zachariah Cherian. 2011. *Rebel Rulers: Insurgent Governance and Civilian Life During War*. Ithaca, NY: Cornell University Press.

Office of the Secretary of Health and Human Services. 1979. "The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research." At http://www.hhs.gov/ohrp/humansubjects/guidance/belmont.html (last accessed 7/5/2015).

Pachirat, Timothy. 2009. "The Political in Political Ethnography: Dispatches from the Kill Floor." In *Political Ethnography: What Immersion Contributes to the Study of Power*, edited by Edward Schatz.. Chicago: University of Chicago Press: 143–162.

———. 2011. *Every Twelve Seconds: Industrialized Slaughter and the Politics of Sight*. New Haven, CT: Yale University Press.

———. 2015. "The Tyranny of Light." *Qualitative and Multi-Method Research: Newsletter of the American Political Science Association's QMMR Section* vol. 13, no. 1.

Parkinson, Sarah Elizabeth. 2013a. "Organizing Rebellion: Rethinking High-Risk Mobilization and Social Networks in War." *American Political Science Review* vol. 107, no. 3: 418–432.

———. 2013b. "Reinventing the Resistance: Order and Violence among Palestinians in Lebanon." Ph.D. Dissertation, Department of Political Science, The University of Chicago.

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

Schwedler, Jillian. 2006. "The Third Gender: Western Female Researchers in the Middle East." *PS: Political Science & Politics* vol. 39, no. 3: 425–428.

Simmons, Erica S. and Nicholas Rush Smith. 2015. "Comparison and Ethnography: What Each Can Learn from the Other." Unpublished Manuscript. University of Wisconsin, Madison.

Straus, Scott. 2005. *The Order of Genocide: Race, Power, and War in Rwanda*. Ithaca, NY: Cornell University Press.

Thomson, Susan. 2010. "Getting Close to Rwandans since the Genocide: Studying Everyday Life in Highly Politicized Research Settings." *African Studies Review* vol. 53, no. 3: 19–34.

Wedeen, Lisa. 2010. "Reflections on Ethnographic Work in Political Science." *Annual Review of Political Science* vol. 13, no. 1: 255–272.

Wood, Elisabeth Jean. 2003. *Insurgent Collective Action and Civil War in El Salvador*. Cambridge: Cambridge University Press.

———. 2006. "The Ethical Challenges of Field Research in Conflict Zones." *Qualitative Sociology* vol. 29, no. 3: 373–386.

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

The Tyranny of Light

Timothy Pachirat
University of Massachusetts, Amherst

In these dark rooms where I live out empty days
I wander round and round
trying to find the windows.

It will be a great relief when a window opens.
But the windows aren't there to be found—
or at least I can't find them. And perhaps
it's better if I don't find them.
Perhaps the light will prove another tyranny.
Who knows what new things it will expose?

—Constantine Cavafy

"I celebrate opacity, secretiveness, and obstruction!" proclaimed no one, ever, in the social sciences.

As with "love" and "democracy," merely uttering the words transparency and openness generates a Pavlovian stream of linguistically induced serotonin. Who, really, would want to come out on record as a transparency-basher, an openness-hater?

But as with love and democracy, it is the specific details of what is meant by transparency and openness, rather than their undeniable power and appeal as social science ideals, that most matter. This, to me, is the single most important point to be made about the DA-RT[1] initiative that has provoked this QMMR symposium:

DA-RT does not equal transparency, and transparency does not equal DA-RT.

Rather, DA-RT is a particular *instantiation*, and—if its proponents have their way—an increasingly institutionalized and "incentivized"[2] *interpretation* of transparency and openness, one which draws its strength from a specific, and contestable, vision of what political science has been—and, equally important—what it should become.

DA-RT proponents argue that they are simply reinforcing a key universal value—transparency—and that they are not doing so in any way that troubles, challenges, reorders, or

[1] Data Access and Research Transparency.

[2] Translation: rewards and punishments can and will be applied for compliance and noncompliance.

imposes an explicit or implicit hierarchy of worth on the ontological and epistemic diversity of existing research communities and traditions within the discipline. DA-RT, in this view, is a strictly neutral vessel, at its core an a-political or depoliticized set of guidelines which scholars from every research tradition should then take and decide for themselves how to best implement and enforce. To wit:

> "…a critical attribute of DA-RT is that it does not impose a uniform set of standards on political scientists."[3]

> "…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."[4]

In this essay, I advance a reading of DA-RT that seeks to trouble its purported neutrality. In particular, I briefly sketch two intertwined dimensions that I believe deserve closer attention and discussion prior to an enthusiastic embrace of discipline-wide transparency standards. The first is historical and contextual, a kind of time-line of DA-RT's development, a sociological account of the key players involved as well as their motivating logics, and a look at the mechanisms of institutionalization, "incentivization," and enforcement that are currently being deployed to normalize DA-RT across the discipline. The second is ontological and epistemological: the rapid, near-automatic, and all-but-unnoticed collapse of the wonderfully ambiguous categories "research communities" and "research traditions" into the two tired but tenacious proxies of "qualitative research" and "quantitative research," proxies that do more in practice to suffocate than to nurture the generative plurality of ontological, epistemological, and even stylistic and aesthetic modes that constitutes the core strength of our discipline.[5]

It is crucial to understand that, on its proponents' own account, the original motivation for both DA-RT and for the APSA Ethics Guidelines Revisions that authorized the DA-RT committee to do its work derive directly from concerns about replicability in empirical research conducted within positivist logics of inquiry. Specifically, "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."[6] As the dominant DA-RT narrative has it, this emerging crisis of replicability in positivist political science[7] was soon found to also exist, in different registers, for a range of scholars "from different methodological and substantive subfields."[8] Thus, while the DA-RT narrative acknowledges its specific and particular origins in concerns over replication of empirical studies conducted within positivist logics of inquiry, it moves quickly from there to claiming a widespread (discipline-wide?) set of shared concerns about similar problems across other methodological and substantive subfields.

DA-RT proponents further assert that, "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."[9] But the DA-RT statement produced by these political scientists, as well as the list of initial DA-RT journal endorsements, strongly suggests that this purported diversity was almost certainly within-set diversity, diversity of subfields and methodologies deployed *within* positivist logics of inquiry rather than across logics of inquiry that might include, for example, interpretive logics of inquiry.[10] As Appendices A and B of APSA's DA-RT statement and the January 2014 *PS* Symposium's disaggregation into DA-RT "in the Qualitative Tradition"[11] versus DA-RT "in the Quantitative Tradition"[12] format suggests, we once again witness the ways in which the type of information a researcher works with (numbers vs. text) simultaneously obscures and usurps a potentially generative discussion about the underlying logics of inquiry that researchers work within and across.

The existing language used to justify DA-RT's universal applicability across the discipline is particularly illuminating here, especially the strongly worded assertion that "[t]he methodologies political scientists use to reach evidence-based conclusions *all* involve extracting information from the social world, analyzing the resulting data, and reaching a conclusion based on that combination of the evidence and its analysis."[13] Attention to the specificity of language deployed here signals immediately that we are working within a decidedly positivist conception of the world. Most scholars working within an interpretivist logic of inquiry would not be so quick to characterize their evidence-based work as being about the *extraction* of *information* from the social world and the subsequent analy-

---

[3] Lupia and Elman 2014, 20.

[4] Elman and Kapiszewski 2014, 44.

[5] For elaboration on this point, see my (2013) review of Gary Goertz and James Mahoney's *A Tale of Two Cultures*.

[6] Lupia and Elman 2014, 19.

[7] See, for example, "Replication Frustration in Political Science," on the Political Science Replication Blog (https://politicalscience replication.wordpress.com/2013/01/03/replication-frustration-in-political-science/, last accessed 6/27/2015). Or, more dramatically, the unfolding and emerging story of outright fabrication, including fabrication of the replication of a fabrication, surrounding LaCour and Green 2014.

[8] Lupia and Elman 2014, 20.

[9] Lupia and Elman 2014, 20.

[10] As of the writing of this essay, <www.dartstatement.org> notes that thus far 25 journal editors have signed on to the DA-RT statement. But there are notable exceptions. Jeff Isaac of *Perspectives on Politics* wrote a prescient letter that outlines not only why he would not sign the DA-RT statement on behalf of *Perspectives* but also why its adoption as a discipline-wide standard might prove detrimental rather than constructive. Other top disciplinary journals that are currently not signatories to DA-RT include *Polity*, *World Politics*, *Comparative Politics*, *ISQ*, and *Political Theory*.

[11] Elman and Kapiszewski 2014.

[12] Lupia and Alter 2014.

[13] Elman and Kapiszewski 2014, 44, emphasis in the original.

sis of the *data* byproducts of this extraction, but would instead speak about the co-constitution of intersubjective knowledge in collaboration with the social world.[14] The D in DA-RT stands, of course, for data, and it is this underlying and unexamined assertion that *all* evidence-based social science is about the *extraction* of *information* which is then subsequently processed and analyzed as *data* in order to produce social science *knowledge* that most clearly signals that the diversity of disciplinary interests represented by DA-RT is both less sweeping and less compelling than is claimed by the dominant DA-RT narrative. In short, quite apart from how DA-RT might be implemented and enforced at the disciplinary level, the very ontological framework of *data extraction* that undergirds DA-RT is itself already anything but neutral with regard to other logics of inquiry that have been long-established as valuable approaches to the study of power.

To be fair, some DA-RT proponents did subsequently advance arguments for the benefits of DA-RT in research communities that do not value replication, as well as in research communities that prize context-specific understanding over generalized explanation.[15] But, as justifications for the importance and necessity of DA-RT *for these communities* in the context of concerns that originated out of a crisis of replication in positivist social science, these arguments seem weak and ad hoc; they sit uncomfortably and awkwardly in the broader frame of *data extraction*; and they fail to demonstrate that members of those communities themselves have been experiencing any sort of crisis of openness or transparency that might cause them to advocate for and invite a disciplinary wide solution like DA-RT.

Take, for example, the surprising assertion advanced in an article on the value of "Data Access and Research Transparency in the Qualitative Tradition" that "[a]lthough the details differ across research traditions, 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.*"[16]

For interpretive methods such as interpretive ethnography, the italicized portion of this statement is nonsensical. There is no such central paradox in interpretive ethnography because the very foundations of interpretive ethnography rest on an ontology in which the social world in which the researcher immerses, observes, and participates is already always co-constituted in intersubjective relationship with the researcher. A work of interpretive ethnography that did not seek to centrally discuss the contours of the researcher's engagement with the social world, that did not aim to detail how the researcher generated and deployed the material that constitutes her ethnography, and that did not strive to share that material in richly specific, extraordinarily lush and detailed language would not just fail to persuade a readership of interpretive ethnographers: it would, literally, cease to be recognizable as a work of interpretive ethnography!

Where other modes of research and writing might prize the construction and presentation of a gleaming and flawless edifice, two key criteria for the persuasiveness of an interpretive ethnography are the degree to which the ethnographer leaves up enough of the scaffolding in her finished ethnography to give a thick sense to the reader of how the building was constructed *and* the degree to which the finished ethnography includes enough detailed specificity, enough rich lushness, about the social world(s) she is interpreting that the reader can challenge, provoke, and interrogate the ethnographer's interpretations using the very material she has provided as an inherent part of the ethnographic narrative.[17]

To put it another way, the very elements of transparency and openness—what interpretive ethnographers often refer to as reflexivity and attention to embodiment and positionality—that DA-RT proponents see as *lacking* in deeply contextual qualitative work constitute the very hallmarks of interpretive ethnography as a mode of research, analysis, and writing. What is more, interpretive ethnography prioritizes dimensions that go beyond what is called for by DA-RT, encouraging its practitioners to ask reflexive questions about positionality and power, including ethnographers' positionality and power as embodied researchers interacting with and producing politically and socially legitimated "knowledge" about the social world, and the potential impacts and effects of that embodied interaction and knowledge production.

Indeed, the types of reflexivity valued by interpretive approaches would question the adequacy of how DA-RT conceives of openness and transparency and would seek instead to examine the power relations implied by a model of research

---

[14] There are several sophisticated treatments of this basic point. See, for example, Patrick Jackson's distinction between dualist and monist ontologies; Dvora Yanow's (2014) essay on the philosophical underpinnings of interpretive logics of inquiry; Peregrine Schwartz-Shea's (2014) distinctions between criteria of evaluation for evidence-based research conducted within positivist and interpretivist logics of inquiry; Lisa Wedeen's (2009) outlining of the key characteristics of interpretivist logics of inquiry; Frederic Schaffer's (2015) treatment of concepts and language from an interpretivist perspective; and Lee Ann Fujii (Forthcoming).

[15] The single paragraph in the pro-DA-RT literature that seems to most directly address interpretive logics of inquiry reads: "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 they can accurately relate such claims to a legitimating context. When readers are empowered to make sense of others [sic] arguments in these ways, the more pathways exist for readers to believe and value knowledge claims" (Lupia and Elman 2014, 22). As I argue below, this paragraph offers a partial description of what interpretive scholars already do, not an argument for why DA-RT is needed.

[16] Elman and Kapiszewski 2014, 46, emphasis mine.

[17] For further elaboration on the importance of reflexivity to interpretive ethnography, see Pachirat 2009a. For key rhetorical characteristics of a persuasive ethnography, see Yanow 2009.

in which the researcher's relationship to the research world is *extractive* in nature and in which transparency and openness are prized primarily in the inter-subjective relationships between researchers and other researchers, but not between the researcher and the research world from which he *extracts information* which he then *processes into data for analysis*. For interpretive ethnographers, research is not an extractive industry like mountain top coal mining, deep water oil drilling, or, for that matter, dentistry. Rather, it is an embodied, intersubjective, and inherently relational enterprise in which close attention to the power relations between an embodied researcher and the research world(s) she moves among and within constitutes a key and necessary part of the interpretive analysis.[18]

In any potential application to interpretive ethnographic research, then, DA-RT seems very much like a solution in search of a problem. Indeed, in its purported neutrality; in its collapsing of "research communities" into the tired but tenacious prefabricated categories of quantitative and qualitative rather than a deeper and much more generative engagement with the diversity of underlying logics of inquiry in the study of power; and in its enforcement through journal policies and discipline-wide ethics guidelines that are insufficiently attentive to ontology and logic-of-inquiry specific diversities, DA-RT risks becoming a solution that generates problems that did not exist before.

Here is one such potential DA-RT generated problem: the claim, already advanced in "Guidelines for Data Access and Research Transparency for Qualitative Research in Political Science," that ethnographers should—in the absence of countervailing human subjects protections or legal concerns—post to a repository the fieldnotes, diaries, and other personal records written or recorded in the course of their fieldwork.[19] But, really, why stop with requiring ethnographers to post their fieldnotes, diaries, and personal records? Why not also require the ethnographer to wear 24 hour, 360 degree, Visual and Audio Recording Technology (VA-RT) that will be digi-

tally livestreamed to an online data repository and time-stamped against all fieldwork references in the finished ethnography? Would the time-stamped, 24 hour, 360 degree VA-RT then constitute the raw "data" that transparently verifies both the "data" and the ethnographer's interpretation and analysis of those data?[20] VA-RT for DA-RT!

VA-RT dramatizes a mistaken view that the ethnographer's fieldnotes, diaries, and personal records constitute a form of raw "data" that can then be checked against any "analysis" in the finished ethnography. The fallacy underlying the mistaken proposal that ethnographic fieldnotes, diaries, and other personal records should be posted to an online repository derives from at least three places.

The first is an extractive ontology inherent in a view of the research world as a source of informational raw material rather than as a specifically relational and deeply intersubjective enterprise. Fieldnotes, and even VA-RT, will always already contain within them the intersubjective relations and the implicit and explicit interpretations that shape both the substance and the form of the finished ethnographic work.[21] Quite simply, there is no prior non-relational, non-interpretive moment of raw information or data to reference back to. What this means is not only that there is no prior raw "data" to reference back to, but that any attempt to de-personalize and remove identifying information from fieldnotes in order to comply with confidentiality and human subjects concerns will render the fieldnotes themselves unintelligible, something akin to a declassified document in which only prepositions and conjunctions are not blacked out.

Second, fieldnotes, far from being foundational truth-objects upon which the "research product" rests, are themselves texts in need of interpretation. Making them "transparent" in an online repository in no way resolves or obviates the very questions of meaning and interpretation that interpretive scholars strive to address.[22]

And third, neither fieldnotes nor VA-RT offer a safeguard "verification" device regarding the basic veracity of a researcher's claims. The researcher produces both, in the end, and both, in the end, are dependent on the researcher's trustworthiness. For it would not be impossible for a researcher to

---

[18] For an example of this kind of close attention to power in relation to a specific fieldsite, see my ethnography in Pachirat 2011. For an example of this kind of reflexive analysis of power at the disciplinary level, see Oren 2003.

[19] See DA-RT Ad Hoc Committee 2014, 26. The specific wording reads: "The document's contents apply to all qualitative analytic techniques employed to support evidence-based claims, as well as all qualitative source materials [including data from interviews, focus groups, or oral histories; fieldnotes (for instance from participant observation or ethnography); diaries and other personal records.…]" I believe concerned ethnographers need to take issue with the underlying logic of this guideline itself and not simply rely on built-in exemptions for human subjects protections to sidestep an attempt to normalize the posting of fieldnotes to repositories. Also note that the main qualitative data repository created in conjunction with DA-RT, the Qualitative Data Repository (QDR), does not contain a single posting of fieldnotes, diaries, or other personal records from ethnographic fieldwork. Where ethnographers have used the QDR, it is to post already publicly available materials such as YouTube clips of public performances. Further, I am not aware of any ethnographic work within political science, anthropology, or sociology for which fieldnotes have been made available in a repository.

[20] And, in the spirit of discipline-wide neutrality, why not implement the same VA-RT requirements for all field researchers in political science, including interviewers, survey-takers, focus-group leaders, and field experimenters? Indeed, why not require VA-RT for large-N statistical researchers as well, and not only during their analysis of existing datasets but also during the prior construction and coding of those data sets? I hope to write more soon on this thought experiment, which provides a nice inversion of a prior ontology-related thought experiment, the Fieldwork Invisibility Potion (FIP). See Pachirat 2009b.

[21] On the inherently interpretive enterprise of writing fieldnotes, see Emerson, Fretz, and Shaw 2011. In particular, Emerson, Fretz, and Shaw demonstrate how fieldnotes, no matter how descriptive, are already filters through which the ethnographer is attending to certain aspects of the research situation over others.

[22] My thanks to Richard Payne for his keen articulation of this point.

fabricate fieldnotes, nor to stage performances or otherwise alter a VA-RT recording.

The notion of a "data repository," either for ethnographic fieldnotes or for VA-RT, is dangerous—at least for interpretive scholarship—both because it elides the interpretive moments that undergird every research interaction with the research world in favor of a non-relational and anonymized conception of "information" and "data," and because it creates the illusion of a fail-proof safeguard against researcher fabrication where in fact there is none other than the basic trustworthiness of the researcher and her ability to communicate that trustworthiness persuasively to her readers through the scaffolding and specificity of her finished ethnography.

Political scientists do not need VA-RT for DA-RT. Instead, we keenly need—to roughly translate back into the language of my positivist friends—a much better specification of DA-RT's scope conditions. Something like DA-RT may indeed be appropriate for positivist traditions of social inquiry in the discipline. But it does not therefore follow that DA-RT should be applied, at however general or abstracted a level, to research communities and traditions in the discipline which are already constituted in their very identity by keen and sustained attention to how the positionality of the researcher in the research world constitutes not only what she sees, but also how she sees it and gives it meaning. Indeed, interpretivists have long argued that scholars working within all modes of inquiry, and the discipline as a whole, would benefit enormously from a much higher level of reflexivity concerning the underpinnings of our research and our knowledge claims. If broader calls for transparency signal a movement toward greater reflexivity within non-interpretivist traditions of social inquiry in the discipline, they deserve both cautious applause and encouragement to expand their existing notions of transparency and openness in ways that acknowledge and embrace their intersubjective relationships and entanglements with the communities, cultures, and ecosystems in which they conduct research and on which their research sometimes returns to act.

References

DA-RT Ad Hoc Committee. 2014. "Guidelines for Data Access and Research Transparency for Qualitative Research in Political Science, Draft August 7, 2013." *PS: Political Science and Politics* vol. 47, no. 1: 25–37.

Elman, Colin, and Diana Kapiszewski. 2014. "Data Access and Research Transparency in the Qualitative Tradition." *PS: Political Science and Politics* vol. 47, no. 1: 43–47.

Emerson, Robert M., Rachel I. Fretz, and Linda L. Shaw. 2011. *Writing Ethnographic Fieldnotes*, Second Edition. Chicago: University of Chicago.

Fujii, Lee Ann. Forthcoming. *Relational Interviewing: An Interpretive Approach*. New York: Routledge.

Jackson, Patrick Thaddeus. 2011. *The Conduct of Inquiry in International Relations*. New York: Routledge.

Lacour, Michael J., and Donald P. Green. 2014. "When Contact Changes Minds: An Experiment on Transmission of Support for Gay Equality." *Science* vol. 346, no. 6215: 1366–1369.

Lupia, Arthur, and George Alter. 2014. "Data Access and Research Transparency in the Quantitative Tradition." *PS: Political Science and Politics* vol. 47, no. 1: 54–59.

Lupia, Arthur, and Colin Elman. 2014. "Introduction: Openness in Political Science: Data Access and Transparency." *PS: Political Science and Politics* vol. 47, no. 1: 19–23.

Oren, Ido. 2003. *Our Enemies and US: America's Rivalries and the Making of Political Science*. Ithaca: Cornell University Press.

Pachirat, Timothy. 2009a. "The Political in Political Ethnography: Dispatches from the Kill Floor." In *Political Ethnography: What Immersion Contributes to the Study of Power*, edited by Edward Schatz. Chicago: University of Chicago Press: 143–161.

———. 2009b. "Shouts and Murmurs: The Ethnographer's Potion." *Qualitative and Multi-Method Research: Newsletter of the American Political Science Association's QMMR Section* vol. 7, no. 2: 41–44.

———. 2011. *Every Twelve Seconds: Industrialized Slaughter and the Politics of Sight*. New Haven: Yale University Press.

———. 2013. "Review of *A Tale of Two Cultures: Qualitative and Quantitative Research in the Social Sciences*." *Perspectives on Politics* vol. 11, no. 3: 979–981.

Schaeffer, Frederic. 2015. *Elucidating Social Science Concepts: An Interpretivist Guide*. New York: Routledge.

Schwartz-Shea, Peregrine. 2015. "Judging Quality: Evaluative Criteria and Epistemic Communities." In *Interpretation and Method: Empirical Research Methods and the Interpretive Turn*, edited by Dvora Yanow and Peregrine Schwartz-Shea. 2nd edition. New York: M.E. Sharpe: 120–145.

Tsoukas, Haridimos. 1997. "The Tyranny of Light." *Futures* vol. 29, no. 9: 827–843.

Wedeen, Lisa. "Ethnography as Interpretive Enterprise." In *Political Ethnography: What Immersion Contributes to the Study of Power*, edited by Edward Schatz. Chicago: University of Chicago Press: 75–93.

Yanow, Dvora. 2009. "Dear Author, Dear Reader: The Third Hermeneutic in Writing and Reviewing Ethnography." In *Political Ethnography: What Immersion Contributes to the Study of Power*, edited by Edward Schatz. Chicago: University of Chicago Press: 275–302.

Yanow, Dvora. 2015. "Thinking Interpretively: Philosophical Presuppositions and the Human Sciences." In *Interpretation and Method: Empirical Research Methods and the Interpretive Turn*, edited by Dvora Yanow and Peregrine Schwartz-Shea. 2nd edition. New York: M.E. Sharpe: 5–26.

Transparency Across Analytic Approaches

Plain Text? Transparency in Computer-Assisted Text Analysis

David Romney
Harvard University

Brandon M. Stewart
Princeton University

Dustin Tingley
Harvard University

In political science, research using computer-assisted text analysis techniques has exploded in the last fifteen years. This scholarship spans work studying political ideology,[1] congressional speech,[2] representational style,[3] American foreign policy,[4] climate change attitudes,[5] media,[6] Islamic clerics,[7] and treaty making,[8] to name but a few. As these examples illustrate, computer-assisted text analysis—a prime example of mixed-methods research—allows gaining new insights from long-familiar political texts, like parliamentary debates, and altogether enables the analysis to new forms of political communication, such as those happening on social media.

While the new methods greatly facilitate the analysis of many aspects of texts and hence allow for content analysis on an unprecedented scale, they also challenge traditional approaches to research transparency and replication.[9] Specific challenges range from new forms of data pre-processing and cleaning, to terms of service for websites, which may explicitly prohibit the redistribution of their content. The Statement on Data Access and Research Transparency[10] provides only very general guidance regarding the kind of transparency positivist empirical researchers should provide. In this paper, we consider the application of these general guidelines to the specific context of computer-assisted text analysis to suggest what transparency demands of scholars using such methods.

We explore the implications of computer-assisted text analysis for data transparency by tracking the three main stages of a research project involving text as data: (1) acquisition, where the researcher decides what her corpus of texts will consist of; (2) analysis, to obtain inferences about the research question of interest using the texts; and (3) *ex post* access, where the researcher provides the data and/or other information to allow the verification of her results. To be transparent, we must document and account for decisions made at each stage in the research project. Transparency not only plays an essential role in replication[11] but it also helps to communicate the essential procedures of new methods to the broader research community. Thus transparency also plays a didactic role and makes results more interpretable.

Many transparency issues are not unique to text analysis. There are aspects of acquisition (e.g., random selection), analysis (e.g., outlining model assumptions), and access (e.g., providing replication code) that are important regardless of what is being studied and the method used to study it. These general issues, as well as a discussion of issues specific to traditional qualitative textual analysis, are outside of our purview. Instead, we focus here on those issues that are uniquely important for transparency in the context of computer-assisted text analysis.[12]

1. Where It All Begins: Acquisition

Our first step is to obtain the texts to be analyzed, and even this simple task already poses myriad potential transparency issues. Traditionally, quantitative political science has been dominated by a relatively small number of stable and publicly available datasets, such as the American National Election Survey (ANES) or the World Bank Development Indicators (WDI). However, a key attraction of new text methods is that they open up the possibility of exploring diverse new types of data, not all of which are as stable and publicly available as the ANES or WDI. Websites are taken down and their content can change daily; social media websites suspend users regularly. In rare cases, websites prohibit any scraping at all. Researchers should strive to record the weblinks of the pages they scraped (and when they scraped them), so that the data can be verified via the Wayback Machine[13] if necessary and available. This reflects a common theme throughout this piece: Full transparency is difficult to impossible in many situations, but researchers should strive to achieve as much transparency as possible.

[1] Laver, Benoit, and Garry 2003.
[2] Quinn et al. 2010.
[3] Grimmer 2010.
[4] Milner and Tingley 2015.
[5] Tvinnereim and Fløttum 2015.
[6] Gentzkow and Shapiro 2010.
[7] Nielsen 2013; Lucas etal. 2015.
[8] Spirling 2011.
[9] King 2011.
[10] DA-RT 2014.

[11] King 2011.
[12] See Grimmer and Stewart (2013) for a review of different text analysis methods.
[13] http://archive.org/web/.

Sometimes researchers are lucky enough to obtain their data from someone who has put it in a tractable format, usually through the platform of a text analytics company. However, this comes with its own problems—particularly the fact that these data often come with severe restrictions on access to the actual texts themselves and to the models used for analysis, making validation difficult. For this reason, some recommend against using black-box commercial tools that do not provide access to the texts and/or the methods used to analyze them.[14] Nevertheless, commercial platforms can provide a valuable source of data that would be too costly for an individual researcher to collect. For instance, Jamal et al.[15] use Crimson Hexagon, a social media analytics company, to look at historical data from Arabic Twitter that would be difficult and expensive to obtain in other ways.[16] In situations where researchers do obtain their data from such a source, they should clearly outline the restrictions placed on them by their partner as well as document how the text could be obtained by another person. For example, in the supplementary materials to Jamal et al.,[17] the authors provide extensive detail on the keywords and date ranges used to create their sample.

A final set of concerns at the acquisition stage is rooted in the fact that, even if texts are taken from a "universe" of cases, that universe is often delimited by a set of keywords that the researcher determined, for example when using Twitter's Streaming API. Determining an appropriate set of keywords is a significant task. First, there are issues with making sure you are capturing all instances of a given word. In English this is hard enough, but in other languages it can be amazingly complicated—something researchers not fluent in the languages they are collecting should keep in mind. For instance, in languages like Arabic and Hebrew you have to take into account gender; plurality; attached articles, prepositions, conjunctions, and other attached words; and alternative spellings for the same word. More than 250 iterations of the Arabic word for "America" were used by Jamal et al.[18] And this number does not take into account the synonyms (e.g., "The United States"), metonyms (e.g., "Washington"), and other associated words that ought to be selected on as well. Computer-assisted selection of keywords offers one potential solution to this problem. For instance, King, Lam, and Roberts[19] have developed an algorithm to help select keywords and demonstrated that it works on English and Chinese data. Such an approach would help supplement researchers' efforts to provide a "universe" of texts related to a particular topic, to protect them from making *ad hoc* keyword selections.

Thus a commitment to transparency at the acquisition stage requires that the researcher either provide the texts they are analyzing, or provide the capacity for the text corpus to be reconstructed. Even when the texts themselves can be made available it is incumbent on the researcher to describe how the collection of texts was defined so that readers understand the universe of texts being considered. When such information is missing it can be difficult for readers and reviewers alike to assess the inferences we can draw from the findings presented.

2. The Rubber Hits the Road: Training and Analysis

Researchers also should be transparent about the decisions made once the data have been collected. We discuss three areas where important decisions are made: text processing, selection of analysis method, and addressing uncertainty.

Processing

All research involves data cleaning, but research with unstructured data involves more than most. The typical text analysis workflow involves several steps at which the texts are filtered and cleaned to prepare them for computer-assisted analysis. This involves a number of seemingly innocuous decisions that can be important.[20]Among the most important questions to ask are: Did you stem (i.e., map words referring to the same concept to a single root)? Which stemmer did you use? If you scraped your data from a website, did you remove all html tags? Did you remove punctuation and common words (i.e. "stop" words), and did you prune off words that only appear in a few texts? Did you include *bigrams* (word pairs) in your analysis? Although this is a long list of items to consider, each is important. For example, removing common stop words can obscure politically interesting content, such as the role of gendered pronouns like 'her' in debates on abortion.[21] Since inferences can be susceptible to these processing decisions, their documentation, in the paper itself or in an appendix, is essential to replication.

Analysis

Once the texts are acquired and processed, the researcher must choose a method of analysis that matches her research objective. A common goal is to assign documents to a set of categories. There are broadly speaking three approaches available: keyword methods, where categories are based on counts of words in a pre-defined dictionary; supervised methods, where humans classify a set of documents by hand (called the training set) to teach the algorithm how to classify the rest; and unsupervised methods, where the model simultaneously estimates a set of categories and assign texts to them.[22]An important part of transparency is justifying the use of a particular approach and clarifying how the method provides leverage to answer the research question. Each method then in turn entails particular transparency considerations.

---

[14] Grimmer and Stewart 2013, 5.

[15] Jamal et al. 2015.

[16] Jamal et al. 2015. In this case, access to the Crimson Hexagon platform was made possible through the Social Impact Program, which works with academics and non-profits to provide access to the platform. Other work in political science has used this data source, notably King, Pan, and Roberts 2013.

[17] Jamal et al. 2015.

[18] Jamal et al. 2015.

[19] King, Lam, and Roberts 2014.

[20] See Lucas et al. (2015) for additional discussion.

[21] Monroe, Colaresi and Quinn 2008, 378.

[22] Grimmer and Stewart 2013, Figure 1.

In supervised learning and dictionary approaches, the best way to promote transparency is to maintain, and subsequently make available, a clear codebook that documents the procedures for how humans classified the set of documents or words.[23] Some useful questions to consider are:

How did the researcher determine the words used to define categories in a dictionary approach?

How are the categories defined and what efforts were taken to ensure inter-coder reliability?

Was your training set selected randomly?

We offer two specific recommendations for supervised and dictionary methods. First, make publicly available a codebook which documents category definitions and the methods used to ensure intercoder reliability.[24] Second, the researcher should report the method used to select the training set of documents—ideally random selection unless there is a compelling reason to select training texts based on certain criteria.[25]

When using unsupervised methods, researchers need to provide justifications for a different set of decisions. For example, such models typically require that the analyst specify the number of topics to be estimated. It is important to note that there is not necessarily a "right" answer here. Having more topics enables a more granular view of the data, but this might not always be appropriate for what the analyst is interested in. The appendix provides one such example using data from political blogs. Furthermore, as discussed by Roberts, Stewart, and Tingley,[26] topic models may have multiple solutions even for a fixed number of topics.[27] These authors discuss a range of stability checks and numerical initialization options that can be employed, which enable greater transparency.

Uncertainty

The final area of concern is transparency in the appropriate incorporation of uncertainty into our analysis. Text analysis is often used as a measurement instrument, with the estimated categorizations used in a separate regression model to, for example, estimate a causal effect. While these two stage models (measure, then regress) are attractive in simplicity, they often come with no straightforward way to incorporate measurement uncertainty. The Structural Topic Model (STM),[28] building on previous work by Treier and Jackman,[29] provides one approach to explicitly building-in estimation uncertainty. But regardless of how estimation uncertainty is modeled, in the spirit of transparency, it is important to acknowledge the concern and specify how it has been incorporated in the analysis.

Another form of uncertainty derives from the research process itself. The decisions discussed in this section—decisions about processing the text, determining dictionary words, or determining categories/the number of categories—are often not the product of a single *ex ante* decision. In reality, the process is iterative. New processing procedures are incorporated, new dictionary words or categories are discovered, and more (or fewer) unsupervised topics are chosen based on the results of previous iterations of the model. While many research designs involve iterative analysis, using text as data often involves more iterations than other research designs. Iteration is a necessary step in the development of any text analysis model, and we are not advocating that researchers unyieldingly devote themselves to their initial course of action. However, we argue that researchers should clearly state when the process was iterative and which aspects of it were iterative. This documentation of how choices were made, in combination with a codebook clearly documenting those choices, helps to minimize remaining uncertainty.

3. On the Other Side: Presentation and Data Access

It is at the access stage of a project that we run into text analyses's most common violation of transparency: not providing replication data. A common reason for not providing these data is that doing so would violate the intellectual property rights of the content provider (website, news agency, etc.), although sometimes the legal concerns are more complicated, such as in the case of research on the Wikileaks cables[30] or Jihadist texts.[31] Sometimes other reasons prevent the researcher from providing the data, such as the sensitivity of the texts or ethical concerns for the safety of those who produced them. Researchers who seek to maximize transparency have found a variety of ways around these concerns. Some provide the document term matrix as opposed to the corpus of texts, allowing for a partial replication of the analysis. For example, the replication material for the recent article by Lucas et al.[32] provides the document term matrix for fatwas used in their analy-

---

[23] Usually those following the instructions in such a codebook are either the researchers themselves or skilled RAs; however, we note with interest a new approach proposed by Benoit et al. (2015), where this portion of the research process is completed via crowdsourcing. In situations where this approach is applicable, it can potentially make replication easier.

[24] See Section 5 of Hopkins et al. (2010) for some compact guidance on developing a codebook. A strong applied example of a codebook is found in the appendix of Stewart and Zhukov (2009).

[25] Random selection of training cases is necessary to ensure that the training set is representative of the joint distribution of features and outcomes (Hand 2006, 7–9; Hopkins and King 2010, 234). Occasionally alternate strategies can be necessary in order to maintain efficiency when categories are rare (Taddy 2013; Koehler-Derrick, Nielsen, and Romney 2015); however, these strategies should be explicitly recognized and defended.

[26] Roberts, Stewart, and Tingley 2015.

[27] Topic models, like Latent Dirichlet Allocation (Blei, Ng, and Jordan 2003) and the Structural Topic Model (Roberts, Stewart, and Tingley 2015), are non-convex optimization problems and thus can have local modes.

[28] See Roberts et al. 2014.

[29] Treier and Jackman 2008.

[30] Gill and Spirling 2015.

[31] Nielsen 2013.

[32] Lucas et al. 2015.

sis. One reason is that some of these fatwas are potentially under copyright and reproducing them would infringe on the rights of their owners, whereas using them for text analysis falls under standard definitions of "fair use" and is not an infringement on copyright. Another possible problem is that disseminating the complete text of the jihadist fatwas in the Lucas et al. data set may raise ethical concerns or legal concerns under US anti-terrorism law—releasing document-term matrices avoids this issue as well. Others provide code allowing scholars who seek to replicate the analysis to obtain the same dataset (either through licensing or web scraping).[33] These are perhaps the best available strategies when providing the raw texts is impossible, but they are each at least partially unsatisfying because they raise the cost of engaging with and replicating the work, which in turn decreases transparency. While intellectual property issues can complicate the creation of replication archives, we should still strive to always provide enough information to replicate a study.

There are also other post-analysis transparency concerns, which are comparatively rarely discussed. We cover two of them here. One of them concerns the presentation of the analysis and results. Greater steps could be taken to allow other researchers, including those less familiar with text analysis or without the computing power to fully replicate the analysis, the opportunity to "replicate" the interpretive aspects of the analysis. As an example, the researcher could set up a browser to let people explore the model and read the documents within the corpus, recreating the classification exercise for those who want to assess the results.[34] With a bit of work, this kind of "transparency through visualization" could form a useful transparency tool.

The second presentation issue relates to the unit of analysis at which research conclusions are drawn. Text analysis is, naturally, done at the text level. However, that is not necessarily the level of interest for the project. Those attempting to use Twitter to measure public opinion, for instance, are not interested in the opinions of the Tweets themselves. But when we present category proportions of texts, that is what we are measuring. As a consequence, those who write the most have the loudest "voice" in our results. To address this issue, we recommend that researchers be more transparent about this concern and either come up with a strategy for aggregating up to the level of interest or justify using texts as the level of analysis.

4. Conclusion

Computer-assisted text analysis is quickly becoming an important part of social scientists' toolkit. Thus, we need to think carefully about the implications of these methods for research transparency. In many ways the concerns we raise here reflect general concerns about transparent and replicable research.[35] However, as we have highlighted, text analysis produces a number of idiosyncratic challenges for replication—from legal restrictions on the dissemination of data to the numerous, seemingly minor, text processing decisions that must be made along the way.

There is no substitute for providing all the necessary data and code to fully replicate a study. However, when full replication is just not possible, we should strive to provide as much information as is feasible given the constraints of the data source. Regardless, we strongly encourage the use of detailed and extensive supplemental appendices, which document the procedures used.

Finally, we do want to emphasize the silver lining for transparency. Text analysis of any type requires an interpretive exercise where meaning is ascribed by the analyst to a text. Through validations of document categories, and new developments in visualization, we are hopeful that the interpretive work can be more fully shared with the reader as well. Providing our readers the ability to not only evaluate our data and models, but also to make their own judgments and interpretations, is the fullest realization of research transparency.

Appendix: An Illustration of Topic Granularity

This table seeks to illustrate the need to be transparent about how researchers choose the number of topics in an unsupervised model, discussed in the paper. It shows how estimating a topic model with different numbers of topics unpacks content at different levels of granularity. The table displays the results of a structural topic model run on the poliblog5k data, a 5,000 document sample of political blogs collected during the 2008 U.S. presidential election season. The dataset is contained in the stm R package (http://cran.r-project.org/web/packages/stm/). The model was specified separately with 5 (left column), 20 (middle column), and 100 (right column) topics.[36] Topics across the different models are aligned according to the correlation in the document-topic loadings—for example, document-topic loadings for "energy" and "financial crisis" in the 20-topic model were most closely matched with that of "economics" in the 5-topic model. With each topic in the left and middle columns, we include words that are highly correlated with those topics. While expanding the number of topics would not necessarily change the substantive conclusions of the researcher, the focus does shift in a way that may or may not be appropriate for a given research question.[37]

---

[33] For example, O'Connor, Stewart, and Smith (2013) use a corpus available from the Linguistic Data Consortium (LDC) which licenses large collections of texts. Although the texts cannot themselves be made publicly available, O'Connor, Stewart, and Smith (2013) provide scripts which perform all the necessary operations on the data as provided by the LDC, meaning that any researcher with access to an institutional membership can replicate the analysis.

[34] One example is the stmBrowser package (Freeman et al., 2015); see also the visualization at the following link: http://pages.ucsd.edu/~meroberts/stm-online-example/index.html.

[35] King 2011.

[36] We used the following specification: stm(poliblog5k.docs, poliblog5k.voc, K=5, prevalence=~rating, data=poliblog5k.meta, init.type="Spectral") with package version 1.0.8

[37] More specifically topics are aligned by correlating the topic-

Economics (tax, legisl, billion, compani, econom)	Energy (oil, energi, tax, economi, price)	Jobs
		Off-shore Drilling, Gas
		Taxes
		Recession/Unemployment
		Fuel Pricing
	Financial Crisis (financi, bailout, mortgag, loan, earmark)	Earmarks
		Inflations/Budget
		Bailouts
		Mortgage Crisis
		Unions
Elections (hillari, poll, campaign, obama, voter)	Parties (republican, parti, democrat, conserv, Pelosi)	Parties
		Congressional Leadership
		Elections
		Corruption/Pork
	Congressional Races (franken, rep, coleman, smith, Minnesota)	Minnesota Race
		Oregon Race
		Martin Luther King
	Biden/Lieberman (biden, joe, debat, lieberman, senat)	Lieberman Campaign
		Debate Night
		Obama Transition Team
		Calls/Meetings
		Senate Votes
		Biden as Running Mate
	Primaries (poll, pennsylvania, virginia, percent, margin)	Polls
		Battleground States
	Republican General (palin, mccain, sarah, john, Alaska)	Attack Adds
		Joe the Plumber
		Gibson Interview
		McCain Campaign
		Palin
		Giuliani
	Candidates (hillari, clinton, deleg, primari, Edward)	Clinton
		Republican Primary Field
		DNC/RNC
		Democratic Primary Field

References

Benoit, Kenneth, Drew Conway, Benjamin E. Lauderdale, Michael Laver, and Slava Michaylov. 2015. "Crowd-Sourced Text Analysis: Reproducible and Agile Production of Political Data." *American Political Science Review*, forthcoming (available at http://eprints.lse.ac.uk/62242/, last accessed 7/5/2015).

Blei, David M., Andrew Y. Ng, and Michael I. Jordan. 2003. "Latent Dirichlet Allocation." *The Journal of Machine Learning Research* vol.3 no.4/5: 993–1022.

"Data Access and Research Transparency (DA-RT): A Joint Statement by Political Science Journal Editors." 2014. At http://media.wix.com/ugd/fa8393_da017d3fed824cf587932534c860ea25.pdf (last accessed 7/10/2015).

Freeman, Michael, Jason Chuang, Margaret Roberts, Brandon Stewart, and Dustin Tingley. 2015. *stmBrowser: An R Package for the Structural Topic Model Browser*. https://github.com/mroberts/stmBrowser (last accessed 6/25/2015).

Gentzkow, Matthew, and Jesse M. Shapiro. 2010. "What Drives Media Slant? Evidence from US Daily Newspapers." *Econometrica* vol. 78, no. 1: 35–71.

Gill, Michael, and Arthur Spirling. 2015. "Estimating the Severity of the Wikileaks U.S. Diplomatic Cables Disclosure." *Political Analysis* vol. 23, no. 2: 299–305.

Grimmer, Justin. 2010. "A Bayesian Hierarchical Topic Model for Political Texts: Measuring Expressed Agendas in Senate Press Releases." *Political Analysis* vol. 18, no. 1: 1–35.

Grimmer, Justin, and Brandon M. Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts." *Political Analysis* vol. 21, no. 2: 1–31.

Hand, David J. 2006. "Classifier Technology and the Illusion of Progress." *Statistical Science* vol. 21, no. 1: 1–15.

Hopkins, Daniel and Gary King. 2010. "A Method of Automated Nonparametric Content Analysis for Social Science." *American Journal of Political Science* vol. 54, no. 1: 229–247.

Hopkins, Daniel, Gary King, Matthew Knowles, and Steven Melendez. 2010. "ReadMe: Software for Automated Content Analysis." Online at http://j.mp/1KbkQ9j (last accessed 7/5/2015).

Jamal, Amaney A., Robert O. Keohane, David Romney, and Dustin Tingley. 2015. "Anti-Americanism and Anti-Interventionism in Arabic Twitter Discourses." *Perspectives on Politics* vol. 13, no. 1: 55–73.

King, Gary. 2011. "Ensuring the Data Rich Future of the Social Sciences." *Science* vol.331 no.6018: 719–721.

King, Gary, Jennifer Pan, and Margaret E. Roberts. 2013. "How Censorship in China Allows Government Criticism but Silences Collective Expression." *American Political Science Review* vol.107 no.2: 326–343.

document loadings (theta) and choosing the topic pairings that maximize the correlation. Topics are then annotated using the 5 most probable words under the given topic-word distribution. We assigned the labels (in bold) based on manual inspection of the most probable words. The most probable words were omitted from the 100 topic model due to space constraints.

Contentious Issues (guy, wright, church, media, man)	Online (linktocommentspostcount, postcounttb, thing, guy, think)	Apologies
		Liberal/Conservative Think Tanks
		Media/Press
		Books Writing
		Religious Words
		Internet Sites
	Social Issues (abort, school, children, gay, women)	Stem Cell Research
		Gay Rights
		Education
		Abortion (Religious)
		Abortion (Women's Rights)
		Health Care
		Family
	Hollywood (doesn, film, didn, isn, eastern)	Radio Show
		Talk Shows
		Emotion Words
		Blogging
		Memes ("Messiah", "Maverick")
		Films and Hollywood
		Eating/Drinking
	Obama Controversies (wright, ayer, barack, obama, black)	Obama Fundraising
		Ayer Issues
		Speeches
		Jewish Community Organization
		Wright
	Climate Change/News (warm, publish, global, newspap, stori)	Climate Change
		Newspapers
		Pentagon Stories
		Violence in News
		Environmental Issues
Legal/Torture (court, investig, tortur, justic, attorney)	Torture (legisl, tortur, court, constitut, law)	Bipartisan Legislation
		CIA Torture
		Rule of Law
		FISA Surveillance
		Supreme Court
		Guantanomo
	Presidency (rove, bush, fox, cheney, white)	Fox News/ Rove
		Cheney Vice Presidency
		Websites
		Bush Legacy
		White House
	Voting Issues (immigr, acorn, illeg, union, fraud)	Voter Fraud
		California Gun Laws
		Illegal Immigration
	Blagojevich and Scandals (investig, blagojevich, attorney, depart, staff)	Blagojevich
		Steven
		Jackson
		Lobbying
		Attorney Scandal
		Johnson
Foreign Military (israel, iran, iraqi, troop, Russia)	Middle East (israel, iran, hama, isra, iranian)	Israel
		Iran Nuclear Weapons
		Saddam Bin Laden Link
		Terrorism in Middle East
	Iraq/Afghanistan Wars (iraqi, troop, iraq, afghanistan, pentagon)	Iraqi Factions
		Pakistan/Afghanistan
		Withdrawal from Iraq
		Surge in Iraq
		Veterans
	Foreign Affairs (russia, world, russian, georgia, democracy)	Russia and Georgia
		Nuclear North Korea
		Rice and Foreign Policy
		Opposition Governments
		American Vision

King, Gary, Patrick Lam, and Margaret E. Roberts. 2014. "Computer-assisted Keyword and Document Set Discovery from Unstructured Text." Unpublished manuscript, Harvard University: http://gking.harvard.edu/publications/computer-Assisted-Keyword-And-Document-Set-Discovery-Fromunstructured-Text (last accessed 7/5/2015).

Koehler-Derrick, Gabriel, Richard Nielsen, and David A. Romney. 2015. "The Lies of Others: Conspiracy Theories and Selective Censorship in State Media in the Middle East." Unpublished manuscript, Harvard University; available from the authors on request.

Laver, Michael, Kenneth Benoit, and John Garry. 2003. "Extracting Policy Positions from Political Texts Using Words as Data." *American Political Science Review* vol 97, no. 2: 311–331.

Lucas, Christopher, Richard Nielsen, Margaret Roberts, Brandon Stewart, Alex Storer, and Dustin Tingley. 2015. "Computer Assisted Text Analysis for Comparative Politics." *Political Analysis* vol.23 no.2: 254–277.

Milner, Helen V. and Dustin H. Tingley. 2015. *Sailing the Water's Edge: Domestic Politics and American Foreign Policy*. Princeton University Press.

Monroe, Burt L., Michael P. Colaresi, and Kevin M. Quinn. 2008. "Fightin' Words: Lexical Feature Selection and Evaluation for Identifying the Content of Political Conflict." *Political Analysis* vol. 16, no. 4: 372–403.

Nielsen, Richard. 2013. "The Lonely Jihadist: Weak Networks and the Radicalization of Muslim Clerics." Ph.D. dissertation, Harvard University, Department of Government. (Ann Arbor: ProQuest/ UMI Publication No. 3567018).

O'Connor, Brendan, Brandon M. Stewart and Noah A. Smith. 2013. "Learning to Extract International Relations from Political Context." In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics, Volume 1: Long Papers* (Sofia, Bulgaria): 1094-1104. Online at http://aclanthology.info/events/acl-2013 and http://aclweb.org/anthology/P/P13/P13-1108.pdf (last accessed 7/10/2015).

Quinn, Kevin M., Burt L. Monroe, Michael Colaresi, Michael H. Crespin, and Dragomir R. Radev. 2010. "How to Analyze Political Attention with Minimal Assumptions and Costs." *American Journal of Political Science* vol. 54, no. 1: 209–228.

Roberts, Margaret, Brandon Stewart, and Dustin Tingley. 2015. "Navigating the Local Modes of Big Data: The Case of Topic Models." In *Computational Social Science: Discovery and Prediction*, edited by R. Michael Alvarez. Cambridge: Cambridge University Press, forthcoming.

Roberts, Margaret E., Brandon M. Stewart, Dustin Tingley, Christopher Lucas, Jetson Leder-Luis, Shana Kushner Gadarian, Bethany Albertson, and David G. Rand. 2014. "Structural Topic Models for Open-Ended Survey Responses." *American Journal of Political Science* vol. 58 no. 4: 1064–1082.

Spirling, Arthur. 2011. "US Treaty Making with American Indians: Institutional Change and Relative Power, 1784–1911." *American Journal of Political Science* vol. 56, no. 1: 84–97.

Stewart, Brandon M., and Yuri M Zhukov. 2009. "Use of Force and Civil-Military Relations in Russia: An Automated Content Analysis." *Small Wars & Insurgencies* vol. 20, no. 2: 319–343.

Taddy, Matt. 2013. "Measuring Political Sentiment on Twitter: Factor Optimal Design for Multinomial Inverse Regression." *Technometrics* vol. 55, no. 4: 415–425.

Treier, Shawn, and Simon Jackman. 2008. "Democracy as a Latent Variable." *American Journal of Political Science* vol. 52, no. 1: 201–217.

Tvinnereim, Endre and Kjersti Fløttum. 2015. "Explaining Topic Prevalence in Answers to Open-Ended Survey Questions about Climate Change." Nature Climate Change, forthcoming (available at http://dx.doi.org/ 10.1038/nclimate2663, last accessed 7/5/2015).

Transparency Standards in Qualitative Comparative Analysis

Claudius Wagemann
Goethe University, Frankfurt

Carsten Q. Schneider
Central European University, Budapest

When judging the usefulness of methods, it is not only their technical principles that matter, but also how these principles are then translated into applied practice. No matter how well developed our techniques and methods are, if their usage runs against their spirit, they cannot be what the originally ancient Greek word "method" literally means: a "way towards a goal." Standards of best practice are therefore important components of methodological advancement, if such standards are recognized for what they ought to be: transitory condensations of a shared understanding that are valid until improved.

The more popular a specific method becomes, the greater the need for shared understandings among users. This was our motivation for proposing a "code of good standards" for Qualitative Comparative Analysis (QCA).[1] Due to the transitory nature of any such list, we subsequently provided an update.[2] Transparency is one of the major underlying themes of this list.

QCA is the most formalized and widespread method making use of set-analytic thinking as a fundamental logical basis for qualitative case analysis.[3] The goal consists of the identification of sufficient and necessary conditions for outcomes, and their derivates, namely INUS and SUIN conditions.[4] Almost by default, QCA reveals conjunctural causation (i.e., conditions that do not work on their own, but have to be combined with one another); equifinality (where more than one conjunction produces the outcome in different cases); and asymmetry (where the complement of the phenomenon is explained in

Claudius Wagemann is Professor for Political and Social Sciences and Dean of Studies at Goethe University, Frankfurt. He is online at wagemann@soz.uni-frankfurt.de and http://www.fb03.uni-frankfurt.de/politikwissenschaft/wagemann. Carsten Q. Schneider is Full Professor and Head of the Political Science Department at the Central European University, Budapest. He is online at schneider@ceu.edu and http://people.ceu.edu/carsten-q_schneider.

[1] Schneider and Wagemann 2010.

[2] In Schneider and Wagemann 2012, 275ff.

[3] Goertz and Mahoney 2012, 11.

[4] INUS stands for "Insufficient but Necessary part of a condition which is itself Unnecessary but Sufficient for the result" (Mackie 1965, 246). SUIN stands for "Sufficient but Unnecessary part of a factor that is Insufficient but Necessary for the result" (Mahoney, Kimball, and Koivu 2009, 126).

different ways than the phenomenon itself).[5] The systematic nature of QCA has resulted in a series of applications from different disciplines.[6]

It is easier to think about transparency issues in QCA when we understand QCA both as a systematic variant of comparative case study methods[7] and as a truth table analysis (which is at the core of the standard QCA algorithm). While regarding QCA as a research approach links the transparency discussion to comparative case-study methods in general, looking at QCA from a technical point of view leads to a discussion of transparency *desiderata* at specific phases of the analysis. After discussing the transparency criteria that derive from these two perspectives, we will briefly discuss more recent developments in QCA methods and link them to the transparency debate.

Transparency for QCA as an Approach

Every QCA should respect a set of general transparency rules, such as making available the raw data matrix, the truth table that is derived from that matrix, the solution formula, and central parameters, such as consistency and coverage measures. If this requires more space than is permitted in a standard article format, then there are numerous ways to render this fundamental information available, either upon request or (better) in online appendices.

Going beyond these obvious requirements for reporting the data and the findings from the analysis, QCA, being a comparative case-oriented method, also has to consider the central elements of *any* comparative case study method when it comes to transparency issues. For instance, case selection is an important step in comparative research. As with comparative methods in general, in QCA there should always be an explicit and detailed justification for the (non-)selection of cases. QCA is usually about middle-range theories, and it is therefore of central importance to explicitly define the reference population.[8] More often than not, the cases used in a QCA *are* the entire target population, but this needs to be stated explicitly.

The calibration of sets is crucial for any set-analytic method and needs to be performed in a transparent manner. For each concept involved in the analysis (conditions and outcome), researchers must determine and transparently justify each case's membership in these sets. This requires both a clear conceptual understanding of the meaning of each set and of the relevant empirical information on each case.[9] All subsequent statements and conclusions about the case and the concept depend on the calibration decisions taken, so transparency regarding calibration decisions is particularly crucial.

The need for transparency with regard to calibration also renders the publication of the raw data matrix from which the calibrated data stem very important. The information contained in such a raw data matrix can consist of anything from standardized off-the-shelf data to in-depth case knowledge, content analysis, archival research, or interviews.[10] Whatever transparency standards are in place for these data collection methods also hold for QCA. Ultimately, transparency on set calibration then culminates in reporting the reasons for choosing the location of the qualitative anchors, especially the 0.5 anchor, as the latter establishes the *qualitative* difference between cases that are more in a set vs. those that are more out.

The need to be transparent about calibration also raises an issue that is sometimes (erroneously) portrayed as a special feature of QCA—the "back and forth between ideas and evidence."[11] "Evidence" generated by a truth table analysis (see below) may provoke new "ideas," which then produce new evidence based on another truth table analysis. This is nothing new for qualitative research. In QCA, "moving back and forth between ideas and evidence" might mean adding and/or dropping conditions or cases, not only as robustness tests, but also prior to that as a result of learning from the data. Initial truth table analysis might reveal that a given condition is not part of any solution formula and thus superfluous, or might suggest aggregating previously separate (but similarly working) conditions in macro-conditions. After initial analyses, scholars might also recalibrate sets. Whether we refer to these as iterative processes or as "serendipity,"[12] there is nothing bad nor unusual about updating beliefs and analytic decisions during the research process—as long, of course, as this is not sold to the reader as a deductive theory-testing story. Indeed, "emphasis on process"[13] is a key characteristic of qualitative research. An important challenge of transparency in QCA research is figuring out how scholars can be explicit about the multi-stage process that led to their results without providing a diary-like account. This also ensures replicability of the results.

Transparency in the Truth Table Analysis

At the heart of any QCA is the analysis of a truth table. Such a truth table consists of all the 2^n logically possible combinations (rows) of the n conditions. For each row, researchers test whether the empirical evidence supports the claim that the row is a subset of the outcome. Those rows that pass the test can be considered sufficient for the outcome. The result of this procedure is a truth table in which each row has one outcome value assigned to it: it is either sufficient for the outcome (1) or not (0), or belongs to a logical remainder.[14] Assignment to the logical remainder means that the particular combination of conditions only exists in theory, but that there is no empirically existing case that exhibits the particular combination that defines this row of the table. This truth table is then subjected to

[5] Schneider and Wagemann 2012, 78ff.

[6] Rihoux et al. 2013.

[7] Regarding QCA as an approach, see Berg-Schlosser et al. 2009.

[8] Ragin 2000, 43ff.

[9] On calibration, see Ragin 2008, 71ff; Schneider and Wagemann 2012, 32ff.

[10] Regarding interviews as the source of a raw data matrix, see Basurto and Speer 2012.

[11] Ragin 1987 and 2000.

[12] Schmitter 2008, 264, 280.

[13] Bryman 2012, 402.

[14] Ragin 2008, 124ff; Schneider and Wagemann 2012, 178ff.

logical minimization, usually with the help of appropriate software capable of performing the Quine-McClusky algorithm of logical minimization.

Both the construction of the truth table and its logical minimization require several decisions: most importantly, how consistent the empirical evidence must be for a row to be considered sufficient for the outcome and how logical remainders are treated. Transparency requires that researchers explicitly state what decisions they have taken on these issues and why they have taken them.

The Treatment of Logical Remainders

More specifically, with regard to the treatment of logical remainders, it is important to note that virtually all social science research, whether based on observational or experimental data, confronts the problem of limited diversity. In QCA, this omnipresent phenomenon manifests itself in the form of logical remainder rows and is thus clearly visible to researchers and readers. This, in itself, is already an important, built-in transparency feature of QCA, which sets it apart from many, if not most, other data analysis techniques.

When confronted with limited diversity, researchers have to make decisions about how to treat the so-called logical remainders because these decisions shape the solution formulas obtained, often labeled as conservative (no assumptions on remainders), most parsimonious (all simplifying assumptions), and intermediate solution (only easy counterfactuals).[15] Researchers need to state explicitly which simplifying assumptions are warranted—and why. Not providing this information makes it difficult, if not impossible, for the reader to gauge whether the QCA results are based on difficult,[16] unwarranted, or even untenable assumptions.[17] Untenable assumptions run counter to common sense or logically contradict each other. Transparency also requires that researchers explicitly report which arguments (sometimes called directional expectations) stand behind the assumptions made. We find it important to note that lack of transparency in the use of logical remainders not only runs counter to transparency standards, but also leaves under-used one of QCA's main comparative advantages: the opportunity to make specific decisions about assumptions that have to be made whenever the data at hand are limited in their diversity.

Handling Inconsistent Set Relations

With regard to deciding whether a row is to be considered consistent with a statement of sufficiency, it has now become a widespread transparency practice to report each row's raw consistency threshold and to report where the cut-off is placed that divides sufficient from not-sufficient rows. Yet, the technical ease with which this parameter is obtained carries the risk that researchers may unintentionally hide from the reader further important information that should guide the decision about whether a row ought to be considered sufficient or not: which of the cases that contradict the statement of sufficiency in a given row, are, in fact, true logically contradictory cases.[18] In fuzzy-set QCA, two truth table rows with the same raw consistency scores can differ in the sense that one contains one or more true contradictory cases whereas the other does not. True contradictory cases hold scores in the condition and outcome sets that not only contradict subset relation statements by degree but also by kind. That is, such cases are much stronger evidence *against* a meaningful subset relation. Hiding such important information behind a (low) consistency score not only contradicts the case-oriented nature of QCA, but is also not in line with transparency criteria.

Presenting The Solution Formula

When reporting the solution formula, researchers must report the so-called parameters of fit, consistency and coverage, in order to express how well the solution fits the data and how much of the outcome is explained by it. To this, we add that the transparency and interpretation of QCA results is greatly enhanced if researchers link the formal-logical solution back to cases. This is best done by reporting—for each sufficient term and the overall solution formula—which of the cases are typical and which of them deviate by either contradicting the statement of sufficiency (deviant cases consistency) or by remaining unexplained (deviant cases coverage).[19]

With regard to this latter point, Thomann's recent article provides a good example. Explaining customization of EU policies, she finds four sufficient terms, i.e., combinations of conditions that are sufficient for the outcome of interest. As can be seen in Figure 1, each term explains a different set of cases; some cases are covered, or explained, by more than one term; and some cases contradict the statement of sufficiency, indicated in bold font. Based on this information, it is much easier for the reader to check the plausibility and substantive interpretation of the solution. Transparency would be further increased by indicating the uniquely covered typical cases, the deviant cases coverage, and the individually irrelevant cases.[20]

New Developments

QCA is rapidly developing.[21] This is why standards of good practice need to be updated. While the principal commitment to transparency remains unaltered, the precise content of this general rule needs to be specified as new issues are brought onto the agenda, both by proponents and critics of QCA.

---

[15] For details, see Ragin 2008, 147ff; Schneider and Wagemann 2012, 167ff.

[16] Ragin 2008, 147ff.

[17] Schneider and Wagemann 2012, 198.

[18] Schneider and Wagemann 2012, 127. Schneider and Rohlfing (2013) use the terminology of deviant cases consistency in kind vs. in degree.

[19] See Schneider and Rohlfing 2013.

[20] See Schneider and Rohlfing (2013) for details.

[21] For a recent overview, see Marx et al. 2014.

Figure 1: Solution Formula with Types of Cases[22]

Table 1: Sufficient conditions for extensive customization

Solution	**RESP*SAL*coerc**	**+ RESP*SAL*RES**	**+ sal*VPL*COERC**	**+ RESP*VPO*COERC ➜ CUSTOM**
Single case coverage	AU:a4 UK:d2,6,7,10,12 13, **a4**	AU:d2, 6,7 FR:d1,2,10,a4,5 GE:d2,4,7,10,a4 UK:d2,6,12	FR:d6,7,9,12,13,a1,3, **d4,8** GE:d6,12,13,a1	AU:d1,2,4,6,7 10,12,13 GE:d1,2,4,6,7,10, 12,13,a4,5
Consistency	0.887	0.880	0.826	0.903
Raw coverage	0.207	0.344	0.236	0.379
Unique coverage	0.033	0.048	0.099	0.076

Solution consistency: 0.805, *solution coverage*: 0.757

Notes: **Bold:** contradictory case.
AU = Austria, FR = France, GE = Germany, UK = United Kingdom.
Raw consistency threshold: 0.764. Next highest consistency score: 0.669
1 path omitted due to low empirical relevance (see online appendix B, Table B3).

Performing Robustness Tests

In QCA, truth table analysis raises particularly important issues of robustness. Scholars need to demonstrate that equally plausible analytic decisions would not lead to substantively different results. Reporting the findings of meaningful robustness tests is therefore a critical aspect of transparency.

The issue of the robustness of QCA results has been on the agenda for quite a while[23] and is currently receiving increased attention.[24] In order to be meaningful, any robustness tests themselves need to be transparent and "[…] need to stay true to the fundamental principles and nature of set-theoretic methods and thus cannot be a mere copy of robustness tests known to standard quantitative techniques."[25] Elsewhere, we have proposed two specific dimensions on which robustness can be assessed for QCA, namely the question of (1.) whether "different choices lead to differences in the parameters of fit that are large enough to warrant a meaningfully different interpretation"[26] and (2.) whether results based on different analytical choices are still in a subset relation. Transparency, in our view, requires that users test whether plausible changes in the calibration (e.g., different qualitative anchors or functional forms), in the raw consistency levels, and in the case selection (adding or dropping cases) produce substantively different results.[27] Any publication should contain at least a footnote, or even better, an (online) appendix explicating the effects, if any, of different analytic choices on the results obtained. Such practice is becoming more and more common.[28]

Revealing Model Ambiguity

It has been known for quite a while in the QCA literature that for one and the same truth table, there can be more than one logically equivalent solution formula. Thiem has recently shown that this phenomenon might be more widespread in applied QCA than thought.[29] Most likely, one reason for the underestimation of the extent of the problem is the fact that researchers often do not report model ambiguity. Transparency dictates, though, that researchers report all different logically equivalent solution formulas, especially in light of the fact that there is not (yet) any principled argument based on which one of these solutions should be preferred for substantive interpretation.

Analyzing Necessary Conditions

For a long time, the analysis of necessity has been treated as a dispensable addendum to the analysis of sufficiency, most likely because the latter is the essence of the truth table analysis, while the former can be easily performed on the basis of isolated conditions and their complements. A recently increased focus on the intricacies of analyzing necessary conditions has triggered two *desiderata* for transparency. First, when assessing the empirical relevance of a necessary condition, researchers should not only check whether the necessary condition is bigger than the outcome set (i.e. is it consistent) and how much so (i.e. is it trivial). In addition, researchers need to report how much bigger the necessary condition is *vis-à-vis* its own negation, i.e., how frequently it occurs in the data. Conditions that are found in (almost) all cases are normally trivially necessary. We propose the Relevance of Necessity (RoN) parameter as a straightforward approach to detecting both sources of

[22] Source: Thomann 2015, 12. Reproduced with permission of the author.
[23] Seawright 2005; Skaaning 2011.
[24] E.g., Krogslund et al. 2014.
[25] Schneider and Wagemann 2012, 285.
[26] Schneider and Wagemann 2012, 286.
[27] Schneider and Wagemann 2012, 287ff.
[28] See, for example, Schneider and Makszin 2014 and their online

appendix at http://dx.doi.org/10.7910/DVN/24456 or Ide, who summarizes his robustness tests in Table 2 (2015, 67).)
[29] Thiem 2014.

trivialness of a necessary condition.[30] Transparency requires that one report this parameter for any condition that is postulated as being a necessary condition. Second, if researchers decide to stipulate two or more single conditions as functionally equivalent, with each of them alone not being necessary but only in their logical union, then all other logical unions of conditions that pass the consistency and RoN test must also be reported. This parallels the above-mentioned need to report model ambiguity during the analysis of sufficiency.

Conclusions

In this brief essay, we have discussed and updated our earlier list of requirements for high-quality QCA with regard to both QCA as a specific technique and QCA as a general research approach. We have then added some reflections on other transparency requirements that follow from recent methodological developments.

Current software developments facilitate many of the transparency requirements. Of most promise here are the R packages specific to QCA and set-theoretic methods more generally.[31] Compared to point-and-click software, script-based software environments facilitate transparency and replicability. Specific robustness test routines can be implemented in the relevant packages; simplifying assumptions can be displayed with a simple command; various forms of graphical representations of solutions can easily produced, etc.

Standards for transparency in QCA research, just as other technical standards, are useful, but only if, or as long as, they do not ossify and turn into unreflected rituals. Researchers cannot be left off the hook when it comes to making decisions on unavoidable trade-offs. Transparency might conflict with other important goals, such as the need to protect information sources, readability, and space constraints. None of this is unique to QCA, though. Even if a full implementation of all best practices is often not feasible, we still deem it important that such standards exist.

References

Basurto, Xavier, and Johanna Speer. 2012. "Structuring the Calibration of Qualitative Data as Sets for Qualitative Comparative Analysis (QCA)." *Field Methods* vol. 24, no. 2: 155–174.

Berg-Schlosser, Dirk, Gisele De Meur, Benoit Rihoux, and Charles C Ragin. 2009. "Qualitative Comparative Analysis (QCA) as an Approach." In *Configurational Comparative Methods:Qualitative Comparative Analysis (QCA) and Related Techniques*, edited by Charles C. Ragin. Thousand Oaks/London: Sage, 1–18.

Bryman, Alan. 2012. *Social Research Methods*. Oxford: Oxford University Press.

Dusa, Adrian, and Alrik Thiem. 2014. Qualitative Comparative Analysis. R Package Version 1.1-4. http://cran.r-project.org/package=QCA (last accessed 6/27/2015).

Goertz, Gary, and James Mahoney. 2012. *A Tale of Two Cultures*. Princeton: Princeton University Press.

Ide, Tobias. 2015. "Why Do Conflicts over Scarce Renewable Resources Turn Violent? A Qualitative Comparative Analysis." *Global Environmental Change* vol. 33 no. 1: 61–70.

Krogslund, Chris, Donghyun Danny Choi, and Mathias Poertner. 2014. "Fuzzy Sets on Shaky Ground: Parametric and Specification Sensitivity in fsQCA." *Political Analysis* vol. 23, no. 1: 21–41.

Marx, Axel, Benoît Rihoux, and Charles Ragin. 2014. "The Origins, Development, and Application of Qualitative Comparative Analysis: The First 25 Years." *European Political Science Review* vol. 6, no. 3: 115–142.

Quaranta, Mario, and Juraj Medzihorsky. 2014. SetMethods. R Package Version 1.1. https://github.com/jmedzihorsky/SetMethods (last accessed 6/27/2015).

Ragin, Charles C. 1987. *The Comparative Method*. Berkeley: The University of California Press.

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

Ragin, Charles C. 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago: University of Chicago Press.

Rihoux, Benoît, Priscilla Álamos-Concha, Daniel Bol, Axel Marx und Ilona Rezsöhazy, 2013. "From Niche to Mainstream Method? A Comprehensive Mapping of QCA Applications in Journal Articles from 1984 to 2011." *Political Research Quarterly* vol. 66, no. 1: 175–185.

Schmitter, Philippe C. 2008. "The Design of Social and Political Research." In *Approaches and Methodologies in the Social Sciences*, edited by Donatella della Porta and Michael Keating. Cambridge: Cambridge University Press, 263–295.

Schneider, Carsten Q., and Kristin Makszin. 2014. "Forms of Welfare Capitalism and Education-Based Participatory Inequality." *Socio-Economic Review* vol. 12, no. 2: 437–462.

Schneider, Carsten Q., and Ingo Rohlfing. 2013. "Set-Theoretic Methods and Process Tracing in Multi-Method Research: Principles of Case Selection after QCA." *Sociological Methods and Research* vol. 42, no. 4: 559–597.

Schneider, Carsten Q., and Claudius Wagemann. 2010. "Standards of Good Practice in Qualitative Comparative Analysis (QCA) and Fuzzy Sets." *Comparative Sociology* vol. 9, no. 3: 397–418.

Schneider, Carsten Q., and Claudius Wagemann. 2012. *Set-Theoretic Methods for the Social Sciences: A Guide for Qualitative Comparative Analysis and Fuzzy Sets in Social Science*. Cambridge: Cambridge University Press.

Seawright, Jason. 2005. "Qualitative Comparative Analysis vis-à-vis Regression." *Studies in International Comparative Development* vol. 40, no. 1: 3–26.

Skaaning, Svend-Erik. 2011. "Assessing the Robustness of Crisp-Set and Fuzzy-Set QCA Results." *Sociological Methods & Research* vol. 40 no. 2: 391–408.

Thiem, Alrik. 2014. "Navigating the Complexities of Qualitative Comparative Analysis: Case Numbers, Necessity Relations, and Model Ambiguities." *Evaluation Review* vol. 38, no. 6: 487–513.

Thomann, Eva. 2015. "Customizing Europe: Transposition as Bottom-up Implementation." *Journal of European Public Policy*, online first, http://dx.doi.org/10.1080/13501763.2015.1008554.

---

[30] Schneider and Wagemann 2012, 220ff.
[31] Dusa and Thiem 2014; Quaranta and Medzihorsky 2014.

Hermeneutics and the Question of Transparency

Andrew Davison
Vassar College

The call for "evidentiary and logical" data-transparency standards[1] in the APSA's "Statement on Data Access and Research Transparency (DA-RT)" opens an important conversation about transparency in political inquiry. In what follows, I consider what transparency might mean in the context of the hermeneutical interpretation of texts, as an empirical and non-positivist mode of political inquiry.

Hermeneutics is itself a diverse tradition. I write primarily within the tradition of the philosophical hermeneutics of Hans-Georg Gadamer, whose insights I have sought to elaborate and illustrate for political explanation. Gadamerian hermeneutics offers what may be understood as guides for transparency. Yet understanding these guides requires delineating how hermeneutics constitutes the explanatory situation of political inquiry in terms very much at odds with the monolingual, empiricist-positivist, data- and method-driven emphases of the DA-RT Statement's vision of transparency. A monolanguage "tends ...to reduce language to the One, that is, to the hegemony of the homogenous."[2] Hermeneutics seriously challenges the idea that there can ever be a single, operational language for constituting anything; in the context of the DA-RT Statement, it parts with the idea that the Statement's empiricist-positivist, social scientific constitutive terms adequately characterize all forms of political inquiry. A brief elaboration of how DA-RT homogenizes all political inquiry in empiricist-positivist terms and thus erases important, alternative empirical and non-scientistic ways of constituting political inquiry seems helpful to clarifying a different, contrasting vision of transparency available within hermeneutical political inquiry.

Essentially, the DA-RT Statement and its supporting literature assume that all political inquirers work in a world of inquiry that understands itself as a "science" of "knowledge" "production" and "transfer," in which researchers "gather," "extract," "collect," and "base conclusions" on "data" (or "evidence"/"information"), "using" different "methods." Indeed, the promulgation of the guidelines follows concerns in the APSA's Governing Council about problems of non-replicability and "multiple instances" in which scholars were unwilling or unable to "provide information" on "how they had derived a *particular conclusion from a specific set of data* or observations."[3] Making "data" available on which "inference and interpretations are based" ("data access"), explaining the gen-

esis of the data ("production transparency"), and demonstrating the data's use ("analytical transparency") are, for the supporters of DA-RT, indispensable for valid, rigorous, accurate, credible, and legitimate knowledge.[4]

The homogenization of political inquiry into data- and method-driven empiricism-positivism is happening in the context of otherwise welcome efforts to open the tent of political inquiry as wide as possible. Lupia and Elman underscore that DA-RT aims for "epistemically neutral" standards that respect the integrity of, apply across, and facilitate communication between diverse Political Science "research traditions" and "communities."[5]

> Political Science is a methodologically diverse discipline, and we are sometimes unable to appreciate how other social scientists generate their conclusions....Higher standards of data access and research transparency will make cross-border understanding more attainable.[6]

But the tent stretches only as far as the terms of data- and method-driven analysis reach. Non-quantitative research is generally categorized as "qualitative" and uniformly characterized in empiricist-positivist terms: "Transparency in any research tradition—whether qualitative or quantitative—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." "[Qualitative] research generally entails … producing thick, rich, and open-ended data;" "Qualitative scholars … gather their own data, rather than rely solely on a shared dataset;" They may not "share a commitment to replication [but] should value greater visibility of data and methods."[7] For Lupia and Elman, "our prescriptive methodologies"—they mention statistical analysis, mathematical modeling, seeking meaning in texts, ethnographic fieldwork, and laboratory experimentation—"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."[8]

There are of course non-quantitative researchers who embrace qualitative empiricist-positivist practice—who have, to emphasize the point, "methodologies" that "extract" "information," "analyze" "resulting data," and "reach" "conclusion[s] based on a combination of evidence and its analysis"—but not all of us engaged in political analysis do, and those of us who constitute our practices in hermeneutical terms do so from principled commitments other than those of empiricism-positivism, with its foundation in the unity of science, and its consequent commitment to methodologically governed sense-data observation, generalizable theorization, and deductive-nomological explanation. The idea of the unity of science suggests that the differences between the non-human natural (usually

---

Andrew Davison is Professor of Political Science at Vassar College. He can be found online at andavison@vassar.edu and at https://politicalscience.vassar.edu/bios/andavison.html.

[1] Lupia and Elman 2014, 20.
[2] Derrida 1998, 40.
[3] Lupia and Elman 2014, 20, emphasis added.

[4] DA-RT Ad Hoc Committee 2014, 25.
[5] Lupia and Elman 2014, 20–22.
[6] Lupia and Elman 2014, 22.
[7] DA-RT Ad Hoc Committee 2014, 27–28; see also Elman and Kapiszewski 2014.
[8] Lupia and Elman 2014, 20 and 22.

empiricist-positivist) sciences and the human-social sciences are a matter of degree—not fundamental—such that the governing terms and practices of the former, like "data and method," may be seamlessly replicated within the latter. The hermeneutic tradition begins elsewhere, viewing the differences between the two domains as fundamental, and, therefore, endorsing different explanatory terms and practices.

While this is not the place for an exhaustive account of Gadamerian hermeneutics, a delineation of its central presuppositions, along with some brief illustrations, may prove helpful for showing what transparency might mean for this research approach. For hermeneutics, social and political explanation presupposes that human beings are meaning-making creatures who constitute their actions, practices, relations, and institutions with purposes that must be accounted for in any compelling explanation. Rather than the scientific theorization of causality through stabilized operational terms, hermeneutics favors the analysis of the multiple and layered, subjective and inter-subjective meanings that constitute texts or text analogues (actions, practices, relations, and institutions). "Constitute" here means to mark the identity of, and distinguish from, other texts or text-analogues that appear similar based on the observation of sense-data alone—such that, where constitutive meanings differ, the text or text-analogues differ as well. Constitutive meanings are, in principle, accessible in and through open-ended actual or metaphorical (as in archival) conversation out of which the interpreter produces accounts or constellations of meanings. These constellations of meanings include both the meaningful horizons of the interpreter that have come to expression through the foregrounding—the bringing to conscious awareness and conversation—of the interpreter's prejudgments in relation to a perplexing question, and the horizons or meanings of the text or text-analogue that the interpreter receives in conversation.

Gadamer referred to this constellation as a *fusion of horizons* (often mistaken for uncritical agreement or affirmation).[9] The fusion entails coming to understand differently: an alteration of some kind in the content of the interpreter's fore-horizons or pre-conversational understanding. "It is enough to say that we understand in a different way, if we understand at all."[10] Understanding in a different way can mean many things. The interpreter—and thus those with whom the interpreter is in explanatory exchange—may come to see the perplexing text or text analogue in either a slightly or radically different way, with a new set of critical concerns, or with a new awareness of the even more opaque quality of the initial perplexity (etc.). The key analytical guideline is to make and keep a conversation going in which some alteration in the interpreter's prior understanding occurs—so that one understands differently (perhaps so differently that while one may disagree with the constitutive meanings of a text, one's disagreement has become nearly impossible, not because one wishes to agree but because one has come to understand).

There is more to say, but already it is possible to identify

___
[9] Gadamer 1989.
[10] Gadamer 1989, 297.

aspects of hermeneutical explanation that may be made transparent, that is, that may be made explicit for other interlocutors as part of conversational inquiry: the identification and expression of the initial perplexity or question that prompted the conversational engagement; the horizons of which the interpreter has become conscious and that have been foregrounded in ongoing conversation; the meanings the interpreter has received in conversation in their multiplicity and constitutively layered qualities; and how, having been in conversation, the interpreter has come to understand differently. The hermeneutical approach therefore constitutes the explanatory situation very differently than data and method-driven, empiricism-positivism.

The constitutive terms are not data and method, but constitutive meaning and conversation (or dialogue, or multilogue). There are no initial data; there are questions or perplexing ideas that reach an interpreter and that the interpreter receives because they have engaged curiosities or questions within the interpreter's fore-horizons. Hermeneutics is empirical (based on the conceptually receptive observation of constitutive meanings) without being empiricist (based on sensory observation, or instrumental sensory enhancement). Meaningfulness—not evidence—arrives, reaches one like a wave; prejudgments occur and reoccur. Hermeneutics envisions foregrounding and putting at play these prejudgments as inquiry. Putting at play means foregrounding for conversation the elements of the interpreter's fore-horizons of which the interpreter becomes, through conversation, conscious. Interpretation involves not collecting or extracting information, but resolute and considerate, conversational receptivity, an openness to receiving meaningfulness from the text—even where the perplexity is surprising or unsettling, and often especially then—and from within the conscious prejudgments that flash and emerge in the reception of the texts. The hermeneutical, conversational interpretation of texts offers access to the concepts meaningfully constitutive of the lives, practices, and relations that inquirers seek to understand or explain.

In conversation, the interpreter aims to let the great variety of happenings of conversation happen: questioning, listening, suggesting, pausing, interrupting, comparing, criticizing, suspending, responding—and additional reflection, suggestion, etc. Conversation does not mean interview or short-lived, harmonious exchange; it entails patient exploration of crucial constitutive meanings. Criticism—to take the most controversial aspect—happens in conversation. (One could say that, in some interlocutive contexts, interlocutors make their strongest assertions in conversation; without such foregrounding there would be no conversation.) But conversation is neither a method nor a strategy; it is a rule of thumb that suggests an interpreter ought to proceed in ways that cannot be anticipated or stipulated in advance, to receive unanticipated, perplexing, novel, or even unsurprising qualities of a text. Not surveying but being led elsewhere in and through actual or metaphorical, reciprocal reflection—receiving/listening, rereceiving/relistening.

Hermeneutical engagement is, therefore, less about reaching a conclusion than immersion in the meaningfulness of difficult conversation. It seeks not to dismiss or to affirm a prior theoretical hunch or conceptual framework (though a new dismissal or *re*-affirmation may constitute a different understanding). Almost without seeking, it seeks a conversational alteration in the interpreter's fore-horizons, and those with whom the interpreter, as *theorist*, converses. As such, hermeneutics requires not concluding but *losing*—not in the sense of not winning, but in the sense of relinquishing or resituating a prior outlook, of allowing something else to happen in one's understanding. Hermeneutics encourages losing in the sense of altering the constitutive horizons of conversationally engaged practices of social and political theorization.

There is more to say abstractly about transparency within hermeneutics, but two all-too brief examples of what losing might mean in the context of explanatory practice may be helpful. Both are drawn from studies concerning the question of political modernity in Turkey, and both involve hermeneutical engagement with texts that express meanings constitutive of political practice. In this regard, these examples also help illustrate how the hermeneutical interpretation of texts relates to the explanation of political practices and power relations. In my studies of what is commonly described as "the secular state" in Turkey, I have suggested that to explain the constitutive vision and practices associated with "Turkey's secular state," inquirers must lose "secularism" for "laicism."[11] "Secularism" meaningfully connotes structural differentiation and separation of spheres and is a central analytical concept in the forehorizon of non-hermeneutical, aspirationally nomological, comparative-empiricist political analysis. "Laicism," derived in Turkey from the French *laïcisme* and brought to expression in Turkish as *laiklik*, connotes possible structural and interpretive integration. Laicism was, moreover and fundamentally, the expressed aim and meaningfully constitutive principle of the founding Kemalist reconfiguration in the 1920s of the prior Ottoman state-Islam relation—a reconfiguration that entailed aspects of separation within overarching constitutive emphases and practices of structural and interpretive integration, control, and official promotion. Hermeneutically speaking, authoritatively describing the founding power relations between state and religion in Turkey as "secular" (with its non- or sometimes anti-religious connotations) thus misses their definitive, constitutive laicist and laicizing (not secular or secularizing) qualities.

It further leads to uncompelling accounts of recent theopolitics in Turkey as radical departure, as opposed to consistent and serious intensification.[12]

A second example: In Mehmet Döºemici's hermeneutical analysis of the mid-twentieth century debate in Turkey over Turkey's application to the European Economic Community, Döºemici effectively suggests that to explain Turkey-Europe relations, inquirers must lose "Turks as not- or not-yet-Euro-

pean" for "Turks as fully European."[13] Hermeneutically engaging texts of profound existential debate between 1959 (Turkey's EEC application) and 1980 (when a military coup squelched the debate), Döºemici illuminates something—modernity in Turkey—that defies empiricist social scientific judgment about Turkey-Europe relations: Between 1959 and 1980, "Turks inquired into who they were and where they were going…. To the extent that this active, self-reflexive and self-defining experience of modernity is historically of European origin…Turkey had during these years, become fully European."[14] Just as to explain the power relations constitutive of Turkey's state, inquirers must lose secularism for laicism, to explain the Turkey-EEC relation as that relation was constituted—*made what it was in Turkey*—inquirers must lose "not/not yet modern/European" for "fully modern/European." Losing as social theoretical alteration—establishing momentarily a new horizon for continuing conversation and thought about secularity, *laïcité*, modernity, Europe, Turkey, Islam, East, West, borders, etc.—happens through hermeneutical engagement with the texts (archives, debates, etc.) expressive of the meanings constitutive of political practice.

In such inquiry, replication of results is not impossible—one may read the texts of the aforementioned analyses and receive the same meaningfulness—but, for credible and legitimate conversation, is also not necessarily desired. There is awareness that any interpreter may see a text (or text-analogue) differently and thus may initiate and make conversation happen from within a different fore-horizon. The range of different interpretations may continually shift; conversation may always begin again and lead somewhere else. There is no final interpretation, not only because others see something entirely new but also because old questions palpitate differently. There is no closing the gates of interpretation around even the most settled question.

The reasons for this open-endedness relate in part to the interplay between subjective and intersubjective meanings, which has implications for transparency. Subjective meanings are purposes constitutive of individual action; intersubjective meanings are shared and contested understandings constitutive of relational endeavors (practices, etc.). For example, a subjective meaning constitutive of my work on *laiklik* is to understand possibilities for organizing the relationship between power and valued traditions; intersubjectively, this meaningfulness is itself constituted by my participation in inquiry (a practice) concerning the politics of secularism. Intersubjectively, none of the terms of my subjective meaning—"understand," "power," "secular," etc.—are "mine." They indicate my participation in a shared, and contested, language of inquiry. Subjectivity is always already intersubjectively constituted, and a full awareness of constitutive intersubjective content eludes the grasp of any interlocutor. Indeed, after Wittgenstein, Marx, Freud, Foucault, and Derrida, there is a compelling recognition that the linguistic, material, psychological, discur-

---

[11] See Davison 1998; Davison 2003.
[12] Parla and Davison 2008.

[13] Döºemici 2013, 227.
[14] Döºemici 2013, 227.

sive, and *differánce* sources of meaning lie prior to or outside the conscious apprehension of interpreting subjects.-

While hermeneutics may therefore be significantly transparent about the happenings of conversation, it is also transparent about the impossibility of complete transparency. Aspects of the subjective and intersubjective meanings of laiklik, for example, may be received in the founding archives, but their constitutedness reaches deep into the lives of the main political actors, Ottoman archives, and those of the French Third Republic. Hermeneutical explanation thus always occurs within unsurpassable limits of subjective and intersubjective human comprehension. Something may always elude, remain ambiguous or opaque, and/or come to conversation for understanding differently—even for understanding hermeneutical understanding differently: My view is that, in favoring constitutive alteration and losing over knowledge accumulation, hermeneutics must be open to losing even its own governing characterization of the hermeneutical situation.

Hermeneutically raising the possibility of losing hermeneutics allows me to underscore what is most compelling about Gadamer's hermeneutics, namely its philosophical, not methodological, emphasis: it does not prescribe rules for interpretation *in order to* understand. Rather, it suggests that the interpretation of meaning happens as if it were a conversation—receiving perplexity, ceaselessly foregrounding fore-horizons, letting the meaningfulness of the perplexity come through, losing, understanding differently. In my work, I have tried to adapt this to inquiry with rules of thumb for interpretation, but the disposition needs no rules. The openness of hermeneutics lies in its not being a method. If one views social life as meaningfully constituted, then conversation is what happens when one understands. And in inquiry, this occurs with a variety of methodological (e.g., contextualist, presentist) and political philosophical (e.g., critical theoretical, conservative, etc.) fore-horizons. Rival interpretations of texts in the history of political thought—*The Prince* teaches evil[15] or a flexible disposition[16] or strategic aesthetic perspectival politics[17]—essentially indicate the putting at play of different pre-judgments in the reception of a perplexing text. Even an imposed, imperial reading may be reconstructed in conversational terms—as the foregrounding of fore-horizons. (Imposition can shut down conversation, but, in some interlocutive contexts, it can also stimulate it.)

From somewhere other than method, one may further say that hermeneutical analysis does not require hermeneutics. Hélène Cixous' "pirate reading" of *The Song of Roland*, which she "detested adored," resembles a conversational engagement:[18] She "abandons" "the idea of fidelity" that she had "inherited from my father" and had "valued above all." "I loved Roland and suddenly"—while struggling to see the face of her schoolgirl classmate, Zohra Drif—"I no longer saw any way to love him, I left him." But she could not "give up reading" and *loses* fidelity for "be[ing] touched on all sides." Conversation—"The song has no tears except on one side."—and she "understands" differently:

> I drew *The Song* toward me but who was I, I did several different readings at the same time when I rebelled with the Saracens. ... before the corpse of the one I could understand Roland's pain I could understand the pain of King Malcud before the other corpse, before each corpse the same pain ... . I could understand the color of the song when I saw that the song sees Count Roland as more and more white the others as more and more black and having only teeth that are white I can do nothing than throw the book across the room it falls under the chair. It's quite useless. I am touched on all sides ... I can no longer close my eyes I saw the other killed, all kill the others, all the others of the others kill the others all smother trust and pity, the spine of the gods split without possible recourse, pride and wrongdoing are on all sides. But the very subtle and passionate song pours all portions of tears over the one to whom it has sworn fidelity. All of a sudden I recognize its incense and fusional blandishment. How is evil beautiful, how beautiful evil is, and how seductive is dreadful pride, I am terrified. I have loved evil, pain, hurt, I hate it, all of a sudden I hatedloved it. The song seduced and abandoned me. No, I abandoned myself to the song. There is no greater treachery.[19]

"Rebelled with the Saracens," "the color of the song," "useless," "its incense," "I am terrified," "hateloved," … all original contributions to knowing *The Song of Roland*. Practicing/not practicing hermeneutics, one must be prepared to let, receive, foreground, converse, lose, and understand differently; and, to underscore, to be open to understanding even these terms (conversation, etc.) differently.[20] As noted above, these aspects of conversational inquiry—the identification of the interpretive perplexity; the letting, foregrounding, and reception of constitutive meaningfulness, both within an interpreter's forehorizons and within texts as they are received conversationally within those horizons; and understanding differently as losing and conceptual alteration—provide the basis for a kind of transparency that may be encouraged in conversation with the aspirations of the DA-RT Statement.

The hermeneutical rejection of the unity of science implies an unfortunate binary between data and meaning, especially insofar as both are situated in a common, and contested, project of social and political explanation. Cross-border work between data- and meaning-governed analyses occurs, and one can be interested in both, in various ways. Yet, within the Gadamerian tradition, among other non-empiricist and non-positivist approaches to political inquiry, the distinction has meaning, and the terms constitutive of one world (e.g., data) are not always meaningful in others. Let's be transparent: to speak as the DA-RT Statement does, in data and method terms,

---

[15] Strauss 1952.

[16] Skinner 1981.

[17] Dietz 1986.

[18] Cixous 2009.

[19] Cixous 2009, 65–67.

[20] Davison 2014.

is to speak a very particular language. Political inquiry is multilingual. The customary tendency at the disciplinary-administrative level is for the standardizing terms of empiricism-positivism to dominate conversation and for hermeneutics not to be read with the relevance to explanation that it understands itself as having.

References

Cixous, Hélène. 2009. *So Close*, translated by Peggy Kamuf. Cambridge: Polity Press.

DA-RT Ad Hoc Committee. 2014. "Guidelines for Data Access and Research Transparency for Qualitative Research in Political Science, Draft August 7, 2013." *PS: Political Science and Politics* vol. 47, no. 1: 25–37.

Davison, Andrew. 1998. *Secularism and Revivalism in Turkey: A Hermeneutic Reconsideration*. New Haven: Yale University Press.

———. 2003. "Turkey a 'Secular' State? The Challenge of Description." *The South Atlantic Quarterly* vol. 102, no. 2/3: 333–350.

———. 2014. *Border Thinking on the Edges of the West: Crossing over the Hellespont*. New York: Routledge.

Derrida, Jacques. 1998. *Monolingualism and the Other; or, the Prosthesis of Origin*, translated by Patrick Mensah. Stanford: Stanford University Press.

Dietz, Mary G. 1986. "Trapping the Prince: Machiavelli and the Politics of Deception." *American Political Science Review* vol. 80, no. 3: 777–799.

Dö°emici, Mehmet. 2013. *Debating Turkish Modernity: Civilization, Nationalism, and the EEC*. New York: Cambridge University Press.

Elman, Colin, and Diana Kapiszewski. 2014. "Data Access and Research Transparency in the Qualitative Tradition." *PS: Political Science and Politics* vol. 47, no. 1: 43–47.

Gadamer, Hans-Georg. 1989. *Truth and Method*, translated by Joel Weinsheimer and Donald G. Marshall. New York: Continuum.

Lupia, Arthur, and Colin Elman. 2014. "Introduction: Openness in Political Science: Data Access and Transparency." *PS: Political Science and Politics* vol. 47, no. 1: 19–23.

Parla, Taha, and Andrew Davison. 2008. "Secularism and Laicism in Turkey," in *Secularisms*, edited by Janet R. Jakobsen and Ann Pellegrini. Durham: Duke University Press, 58–75.

Skinner, Quentin. 1981. *Machiavelli*. New York: Hill and Wang.

Strauss, Leo. 1952. *Persecution and the Art of Writing*. Glencoe: The Free Press.

Reflections on Analytic Transparency in Process Tracing Research

Tasha Fairfield
London School of Economics and Political Science

While the aims of APSA's Data Access and Research Transparency (DA-RT) initiative are incontrovertible, it is not yet clear how to best operationalize the task force's recommendations in the context of process tracing research. In this essay, I link the question of how to improve analytic transparency to current debates in the methodological literature on how to establish process tracing as a rigorous analytical tool. There are tremendous gaps between recommendations and actual practice when it comes to improving and elucidating causal inferences and facilitating accumulation of knowledge. In order to narrow these gaps, we need to carefully consider the challenges inherent in these recommendations alongside the potential benefits. We must also take into account feasibility constraints so that we do not inadvertently create strong disincentives for conducting process tracing.

Process tracing would certainly benefit from greater analytic transparency. As others have noted,[1] practitioners do not always clearly present the evidence that substantiates their arguments or adequately explain the reasoning through which they reached casual inferences. These shortcomings can make it very difficult for scholars to interpret and evaluate an author's conclusions. At worst, such narratives may read as little more than potentially plausible hypothetical accounts.

Researchers can make significant strides toward improving analytic transparency and the overall quality of process tracing by (a) showcasing evidence in the main text as much as possible, including quotations from interviews and documents wherever relevant, (b) identifying and discussing background information that plays a central role in how we interpret evidence, (c) illustrating causal mechanisms, (d) assessing salient alternative explanations, and (e) including enough description of context and case details beyond our key pieces of evidence for readers to evaluate additional alternative hypotheses that may not have occurred to the author. Wood's research on democratization from below is a frequently lauded example that illustrates many of these virtues.[2] Wood clearly articulates the causal process through which mobilization by poor and working-class groups led to democratization in El

[1] Elman and Kapiszewski 2014; Moravcsik 2014.
[2] Wood 2000; Wood 2001.

Salvador and South Africa, provides extensive and diverse case evidence to establish each step in the causal process, carefully considers alternative explanations, and explains why they are inconsistent with the evidence. Wood's use of interview evidence is particularly compelling. For example, in the South African case, she provides three extended quotations from business leaders that illustrate the mechanism through which mobilization led economic elites to change their regime preferences in favor of democratization: they came to view democracy as the only way to end the massive economic disruption created by strikes and protests.[3]

Beyond these sensible if potentially demanding steps, can we do more to improve analytic transparency and causal inference in process tracing? Recent methodological literature has suggested two possible approaches: explicit application of Van Evera's (1997) process tracing tests, and the use of Bayesian logic to guide inference. As a practitioner who has experimented with both approaches and compared them to traditional narrative-based process tracing, I would like to share some reflections from my own experience that I hope will contribute to the conversation about the extent to which these approaches may or may not enhance analytic transparency.

I became interested in the issue of analytic transparency after submitting an article manuscript on strategies for taxing economic elites in unequal democracies, which included four Latin American case studies. The case narratives employed process tracing to illustrate the causal impact of reform strategies on the fate of proposed tax-reform initiatives. I marshaled key pieces of evidence from in-depth fieldwork, including interviews, congressional records, and newspaper reports to substantiate my arguments. However, several of the reviews that I received upon initial submission questioned the contribution of qualitative evidence to the article's causal argument. For example, a reviewer who favored large-n, frequentist hypothesis-testing objected that the case evidence was anecdotal and could not establish causality. Another reviewer was skeptical of the hypothesis that presidential appeals invoking widely shared values like equity could create space for reforms that might not otherwise be feasible—a key component of my explanation of how the center-left Lagos administration in Chile was able to eliminate a regressive tax benefit—and felt that the case study did not provide enough evidence to substantiate the argument.

These reviews motivated me to write what I believe is the first published step-by-step account that explicitly illustrates how process-tracing tests underpin inferences drawn in a case narrative.[4] I chose one of the four case studies and systematically identified each piece of evidence in the case narrative. I also included several additional pieces of evidence beyond those present in the text to further substantiate my argument. Applying state-of-the-art methods literature, I explained the logical steps that allowed me to draw causal inferences from each piece of evidence and evaluated how strongly each piece of evidence supported my argument with reference to particular types of tests. For example, I explained that specific statements from opposition politicians indicating that President Lagos' equity appeal compelled the right-party coalition to reluctantly accept the tax reform provide very strong evidence in favor of my explanation, not only because these statements affirm that explanation and illustrate the underlying causal mechanism, but also because it would be extremely surprising to uncover such evidence if the equity appeal had no causal effect.[5] Based on these observations, the equity appeal hypothesis can be said to pass "smoking gun" tests: this evidence is not *necessary* to establish the hypothesis, but it can be taken as *sufficient* to affirm the hypothesis.[6] Stated in slightly different terms, the evidence is not *certain*—hearing right-wing sources confess that the government's strategy forced their hand is not an ineluctable prediction of the equity-appeal hypothesis, but it is *unique* to that hypothesis—these observations would not be predicted if the equity hypothesis were incorrect.[7] Other types of tests (hoop, straw-in-the-wind, doubly decisive) entail different combinations of these criteria.

While this exercise was most immediately an effort to convince scholars from diverse research traditions of the soundness of the article's findings, this type of procedure also advances analytic transparency by helping readers understand and assess the research. Scholars cannot evaluate process tracing if they are not familiar with the method's logic of causal inference, if they are unable to identify the evidence deployed, or if they cannot assess the probative weight of the evidence with respect to the explanation. While I believe that well-written case narratives can effectively convey all of this information to readers who are familiar with process tracing, explicit pedagogical appendices may help make process tracing more accessible and more intelligible for a broad audience.

However, there are drawbacks inherent in the process-tracing tests approach. For example, evidence rarely falls into the extreme categories of necessity and sufficiency that are generally used to classify the four tests. For that reason, I found it difficult to cast inferences in these terms; the pieces of evidence I discussed in my appendix did not all map clearly onto the process-tracing tests typology. Furthermore, it is not clear how the results of multiple process-tracing tests should be aggregated to assess the strength of the overall inferences in cases where the evidence does not line up neatly in favor of a single explanation.

These problems with process-tracing tests motivated me to redo my appendix using Bayesian analysis. This endeavor is part of a cross-disciplinary collaboration that aims to apply insights from Bayesian analysis in physics to advance the growing methodological literature on the Bayesian underpinnings of process tracing.[8] We believe the literature on process-tracing tests has rightly made a major contribution to qualitative methods. Yet Bayesian analysis offers a more pow-

---

[3] Wood 2001, 880.
[4] Fairfield 2013.

[5] See Fairfield 2013, 56 (observations 2a-e).
[6] Collier 2011; Mahoney 2012.
[7] Van Evera 1997; Bennett 2010.
[8] Fairfield and Charman 2015.

erful and more fundamental basis for understanding process tracing. Instead of asking whether a single hypothesis passes or fails a series of tests, which is very close to a frequentist approach, Bayesian analysis asks whether our evidence makes a given hypothesis more or less plausible compared to rivals, taking into account our prior degree of belief in each hypothesis and relevant background information that helps us interpret the evidence. While process-tracing tests can be incorporated within a Bayesian framework as special cases,[9] Bayesian analysis allows us to avoid the restrictive language of necessity and sufficiency by focusing on the degree to which a given piece of evidence alters our confidence in a hypothesis relative to rivals. Moreover, Bayesian probability provides clear procedures for aggregating inferences from distinct pieces of evidence.

Literature on *informal* Bayesianism in process tracing has elucidated various best practices that enhance analytic transparency.[10] One key lesson is that what matters most for inference is not the amount of evidence but rather how decisive the evidence is relative to the hypotheses at hand. In some cases, one or two highly probative pieces of evidence may give us a high level of confidence in an explanation. However, the available evidence does not always allow us to draw definitive conclusions about which hypothesis provides the best explanation, in which case we should openly acknowledge that uncertainty remains, while working hard to obtain more probative evidence where possible.[11]

Recently, several scholars have taken a step further by advocating that Bayesian analysis in process tracing should be *formalized* in order to make causal inferences more systematic, more explicit, and more transparent.[12] By revising my tax-reform appendix with direct applications of Bayes' theorem—the first such exercise of its kind—my collaborator and I aim to illustrate what formalization would entail for qualitative research that draws on extensive case evidence and to assess the advantages and disadvantages of this approach. I begin with an overview of the substantial challenges we encountered and then discuss situations in which formalization might nevertheless play a useful role in advancing analytic transparency.

First, formalizing Bayesian analysis requires assigning numerical values to all probabilities of interest, including our prior degree of belief in each rival hypothesis under consideration and the likelihood of observing each piece of evidence if a given hypothesis is correct. This task is problematic when the data are inherently qualitative. We found that our numerical likelihood assignments required multiple rounds of revision before they became reasonably stable, and there is no guarantee that we would have arrived at similar values had we approached the problem from a different yet equally valid starting point.[13] We view these issues as fundamental problems for

advocates of quantification that cannot easily be resolved either through efforts at standardization of practice or by specifying a range of probabilities rather than a precise value. The latter approach relocates rather than eliminates the arbitrariness of quantification.[14]

Second, highly formalized and fine-grained analysis ironically may obscure rather than clarify causal inference. Disaggregating the analysis to consider evidence piece-by-piece risks compromising the level on which our intuitions can confidently function. In the tax reform case we examine, it strikes us as intuitively obvious that the total body of evidence overwhelmingly favors a single explanation; however, reasoning about the contribution of each piece of evidence to the overall conclusion is much more difficult, all the more so if we are trying to quantify our reasoning. If we disaggregate the evidence too finely and explicitly unpack our analysis into too many steps, we may become lost in minutiae. As such, calls for authors to "detail the micro-connections between their data and claims… and discuss how evidence was aggregated to support claims,"[15] which seem entirely reasonable on their face, could actually lead to less clarity if taken to extremes.

Third, formal Bayesian analysis becomes intractable in practice as we move beyond very simple causal models, which in our view are rarely appropriate for the social sciences. Whereas frequentists consider a single null hypothesis and its negation, applying Bayes' theorem requires elaborating a complete set of mutually exclusive hypotheses. We need to explicitly state the alternatives before we can reason meaningfully about the likelihood of observing the evidence if the author's hypothesis does not hold. Ensuring that alternative hypotheses are mutually exclusive is nontrivial and may entail significant simplification. For example, some of the hypotheses we assess against my original explanation in the revised appendix involve causal mechanisms that—in the real world—could potentially operate in interaction with one another. Assessing such possibilities would require carefully elaborating additional, more complex, yet mutually exclusive hypotheses and would aggravate the challenges of assigning likelihoods to the evidence uncovered. By contrast, in the natural sciences, Bayesian analysis is most often applied to very simple hypothesis spaces (even if the underlying theory and experiments are highly complex); for example: H_1 = the mass of the Higgs boson is between 124 and 126 GeV/c^2, H_2 = the mass falls between 126 and 128 GeV/c^2, and so forth.

---

regardless of the order in which we incorporate each piece of evidence into our analysis. Literature in the subjective Bayesian tradition has sometimes maintained that the order in which the evidence is incorporated does matter, but we view that approach as misguided and that particular conclusion as contrary to the laws of probability. These points are further elaborated in Fairfield and Charman 2015.

[14] In their work on critical junctures, Capoccia and Kelemen (2007, 362) likewise note: "While historical arguments relied on assessments of the likelihood of various outcomes, it is obviously problematic to assign precise probabilities to predictions in historical explanations…."

[15] DA-RT Ad Hoc Committee 2014, 33.

---

[9] Humphreys and Jacobs forthcoming.

[10] Bennett and Checkel 2015.

[11] Bennett and Checkel 2015a, 30f.

[12] Bennett and Checkel 2015b, 267; Bennett 2015, 297; Humphreys and Jacobs forthcoming; Rohlfing 2013.

[13] Bayes' theorem implies that we must reach the same conclusions

Numerous other practical considerations make formal Bayesian analysis infeasible beyond very simple cases. The Chilean tax reform example I chose for the original appendix is a particularly clear-cut case in which a small number of key pieces of evidence establish the causal importance of the reform strategy employed. The original case narrative was 583 words; the more extensive case narrative in my book, *Private Wealth and Public Revenue in Latin America: Business Power and Tax Politics,* is 1,255 words. By comparison, the original process-tracing tests appendix was 1,324 words; our Bayesian version is presently roughly 10,000 words. My book includes 33 additional case studies of tax reform initiatives. If scholars were expected to explicitly disaggregate and elaborate process tracing to the extent that we have done in our Bayesian appendix, it would be a death knell for qualitative research. Hardly anyone would undertake the task; the timeline for producing publishable research—which is already long for case-based qualitative work—would become prohibitive.

To be sure, no one has suggested such stringent standards. Advocates of Bayesian process tracing have been clear that they do not recommend full quantification in all cases. Yet we fear that there may be little productive middle ground between qualitative process tracing underpinned by informal Bayesian reasoning and full quantification in order to apply Bayes' theorem. Attempts to find a middle ground risk disrupting clear and cogent narratives without providing added rigor, since they would not be able to employ the mathematical apparatus of Bayesian probability. We are therefore skeptical of even "minimal" recommendations for scholars to identify their priors and likelihood ratios for the most probative pieces of evidence.[16] The question of how to make process tracing more analytically explicit without risking false precision is an important problem for methodologists and practitioners to grapple with moving forward.

Given these caveats, when might formal Bayesian analysis contribute to improving causal inference and analytic transparency in qualitative research? First and foremost, we see an important pedagogical role. Reading examples and trying one's own hand at such exercises could help to familiarize students and established practitioners with the inferential logic that underpins process tracing. These exercises might also help train our intuition to follow the logic of Bayesian probability more systematically. Bayesianism is much closer than frequentism to how we intuitively reason in the face of uncertainty, but we need to learn to avoid biases and pitfalls that have been well documented by cognitive psychologists. As Bennett notes, "further research is warranted on whether scholars … reach different conclusions when they use Bayesian mathematics explicitly rather than implicitly, and whether explicit use of Bayesianism helps to counteract the cognitive

biases identified in lab experiments."[17] We explore these questions with regard to our own reasoning about the Chilean tax reform case. On the one hand, we have not identified any inferential differences between the original case narrative and the formalization exercise. This consistency could indicate that informal Bayesian reasoning functioned very well in this instance, or that the intuition underpinning that informal analysis also strongly shaped the (necessarily somewhat ad-hoc) quantification process. On the other hand, we do note several differences between our Bayesian analysis and the process-tracing tests approach regarding the inferential weights assigned to distinct pieces of evidence. The lesson is that explicitly elaborating alternative hypotheses, rather than attempting to assess a single hypothesis (the equity appeal had an effect) against its negation (it had no effect), can help us better assess the probative value of our evidence.[18]

Second, these exercises could play a role in elucidating the precise locus of contention when scholars disagree on causal inferences drawn in a particular case study. We explore how this process might work in our revised appendix. We first assign three sets of priors corresponding to different initial probabilities for my equity-appeal hypothesis and three rival hypotheses. For each set of priors, we then calculate posterior probabilities across three scenarios in which we assign relatively larger or smaller likelihood ratios for the evidence. We find that in order to remain unconvinced by my explanation, a skeptical reader would need to have extremely strong priors against the equity-appeal hypothesis and/or contend that the evidence is far less discriminating (in terms of likelihood ratios) than we have argued. While identifying the precise points of disagreement could be inherently valuable for the knowledge accumulation process, formal Bayesian analysis may be less effective for resolving disputes in cases that are less clear-cut than the one we have examined. Scholars may well continue to disagree not only on prior probabilities for hypotheses, but more importantly on the probative weight of key pieces of evidence. Such disagreements may arise from differences in personal judgments as well as decisions about how to translate those judgments into numbers.

Third, elaborating a formal Bayesian appendix for an illustrative case from a scholar's own research might help establish the scholar's process tracing "credentials" and build trust among the academic community in the quality of the scholar's analytical judgments. As much as we try to make our analysis transparent, multiple analytical steps will inevitably remain implicit. Scholars who conduct qualitative research draw on vast amounts of data, often accumulated over multiple years of fieldwork. Scholars often conduct hundreds of interviews, to mention just one type of qualitative data. There is simply too much evidence and too much background information that informs how we evaluate the evidence to fully articulate or catalog. At some level, we must trust that the scholar has made sound judgments along the way; qualitative research is simply not replicable as per a laboratory science desideratum. But of

---

[16] See Bennett and Checkel 2015b, 267. We would further argue that the most probative pieces of evidence are precisely those for which quantification is least likely to provide added value. The author can explain why the evidence is highly decisive without need to invent numbers, and if the evidence is indeed highly decisive, readers should be able to recognize it as such on its face.

[17] Bennett 2015, 297.

[18] For a detailed discussion, see Fairfield and Charman 2015, 17f.

course trust in analytical judgment must be earned by demonstrating competence. Scholars might use a formalized illustration to demonstrate their care in reasoning about the evidence and the plausibility of the assumptions underlying their inferences. Again, however, further research is needed to ascertain whether the ability to formalize improves our skill at informal analysis, and to a significant extent, moreover, the quality of informal process tracing can be assessed without need for quantifying propositions.

To conclude, my experiments with explicit application of process-tracing tests and formal Bayesian analysis have been fascinating learning experiences, and I believe these approaches provide critical methodological grounding for process tracing. Yet I have become aware of limitations that restrict the utility and feasibility of formalization and fine-grained disaggregation of inferences in substantive process tracing. There is certainly plenty of scope to improve analytic transparency in process-tracing narratives—e.g. by highlighting the evidence and explaining the rationale behind nuanced inferences. Future methodological research may also provide more insights on how to make informal Bayesian reasoning more systematic and rigorous without recourse to quantification. Increasing analytic transparency in process tracing, in ways that are feasible for complex hypotheses and extensive qualitative evidence, will surely be a key focus of methodological development in years to come. In the meantime, further discussion about the practical implications of the DA-RT analytic transparency recommendations for qualitative research is merited.

References

Bennett, Andrew, and Jeffrey Checkel, eds. 2015. *Process Tracing in the Social Sciences: From Metaphor to Analytic Tool*. New York: Cambridge University Press.

Bennett, Andrew, and Jeffrey Checkel. 2015a. "Process Tracing: From Philosophical Roots to Best Practices." In *Process Tracing in the Social Sciences: From Metaphor to Analytic Tool*, edited by Andrew Bennett and Jeffrey Checkel. New York: Cambridge University Press, 3–37.

Bennett, Andrew, and Jeffrey Checkel. 2015b. "Beyond Metaphors: Standards, Theory, and the 'Where Next' for Process Tracing." In *Process Tracing in the Social Sciences: From Metaphor to Analytic Tool*, edited by Andrew Bennett and Jeffrey Checkel. New York: Cambridge University Press, 260–275.

Bennett, Andrew. 2015. "Appendix: Disciplining Our Conjectures: Systematizing Process Tracing with Bayesian Analysis." In *Process Tracing in the Social Sciences: From Metaphor to Analytic Tool*, edited by Andrew Bennett and Jeffrey Checkel. New York: Cambridge University Press, 276–298.

———. 2010. "Process Tracing and Causal Inference." In *Rethinking Social Inquiry*, edited by Henry Brady and David Collier. 2nd edition. Lanham: Rowman and Littlefield: 207–220.

Capoccia, Giovanni, and R. Daniel Kelemen. 2007. "The Study of Critical Junctures." *World Politics* vol. 59, no. 2: 341–369.

Collier, David. 2011. "Understanding Process Tracing." *PS: Political Science and Politics* vol. 44, no. 4: 823–830.

DA-RT Ad Hoc Committee. 2014. "Guidelines for Data Access and Research Transparency for Qualitative Research in Political Science, Draft August 7, 2013." *PS: Political Science and Politics* vol. 47, no. 1: 25–37.

Elman, Colin, and Diana Kapiszewski. 2014. "Data Access and Research Transparency in the Qualitative Tradition." *PS: Political Science and Politics* vol. 47, no. 1: 43–47.

Fairfield, Tasha. 2013. "Going Where the Money Is: Strategies for Taxing Economic Elites in Unequal Democracies." *World Development* vol. 47 (July): 42–57.

———. 2015. *Private Wealth and Public Revenue in Latin America: Business Power and Tax Politics*. New York: Cambridge University Press.

Fairfield, Tasha, and Andrew Charman. 2015. "Bayesian Probability: The Logic of (Political) Science." Prepared for the Annual Meeting of the American Political Science Association, Sept. 3-6, San Francisco.

———. 2015. "Applying Formal Bayesian Analysis to Qualitative Case Research: an Empirical Example, Implications, and Caveats." Unpublished manuscript, London School of Economics. (http://eprints.lse.ac.uk/62368/, last accessed July 6, 2015).

Humphreys, Macartan, and Alan Jacobs. Forthcoming. "Mixing Methods: A Bayesian Approach." *American Political Science Review*.

Mahoney, James. 2012. "The Logic of Process Tracing Tests in the Social Sciences." *Sociological Methods and Research* vol. 41, no. 4: 570–597.

Rohlfing, Ingo. 2013. "Bayesian Causal Inference in Process tracing: The Importance of Being Probably Wrong." Paper presented at the Annual Meeting of the American Political Science Association, Aug.29-Sept.1, Chicago. (http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2301453, last accessed 7/5/2015).

Van Evera, Stephen. 1997. *Guide to Methods for Students of Political Science*. Ithaca, NY: Cornell University Press.

Wood, Elisabeth. 2000. *Forging Democracy from Below: Insurgent Transitions in South Africa and El Salvador*. New York: Cambridge University Press.

———. 2001. "An Insurgent Path to Democracy: Popular Mobilization, Economic Interests, and Regime Transition in South Africa and El Salvador." *Comparative Political Studies* vol. 34, no. 8: 862–888.

Conclusion: Research Transparency for a Diverse Discipline

Tim Büthe
Duke University

Alan M. Jacobs
University of British Columbia

The contributors to this symposium offer important reflections and insights on what research transparency can and should mean for political scientists. In offering these insights, they have drawn on their experience and expertise in a broad range of research traditions that prominently involve one or more "qualitative" methods for gathering or analyzing empirical information. The issues discussed in this symposium, however, are just as important for scholars in research traditions that use primarily or exclusively "quantitative" analytical methods and pre-existing datasets, as we will discuss below.

Rather than simply summarize the many important points that each contributor has made, we seek in this concluding essay to map out the conversation that has unfolded in these pages—in particular, to identify important areas of agreement about the meaning of transparency and to illuminate the structure and sources of key disagreements. We also reflect on broader implications of the symposium discussion for the transparency agenda in Political Science.

To organize the first part of our discussion, we largely employ the APSA *Ethics Guide's* distinction among production transparency (defined as providing an "account of the procedures used to collect or generate data"[1]), analytic transparency (defined as "clearly explicating the links connecting data to conclusion"[2]), and data access.

These categories are closely related to matters of empirical evidence, and this will also be our primary focus below. We wish to emphasize at the outset, however, that openness in the research process is not solely a matter of how we account for and share data and data-analytic procedures. Research transparency—defined broadly as providing a full account of the sources and content of ideas and information on which a scholar

has drawn in conducting her research, as well as a clear and explicit account of how she went about the analysis to arrive at the inferences and conclusions presented—begins with the conceptual and theoretical work that must, at least in part, take place prior to any empirical research.

If key concepts are unclear (and not just because they are "essentially contested"[3]), then a scholar's ability to communicate her research question, specify possible answers, and articulate any findings in a reliably comprehensible manner will be severely impaired. Similarly, if references to the work of others on whose ideas we have built are missing, erroneous, or incomplete to the point of making those sources hard to find, we are not only failing to fully acknowledge our intellectual debts, but are also depriving the scholarly community of the full benefit of our work. In his symposium essay, Trachtenberg calls for a revival of the long-standing social scientific norm regarding the use of what legal scholars call "pin-cites"— i.e., citations that specify the particular, pertinent passages or pages rather than just an entire work (whenever a citation is not to a work in its entirety or its central argument/finding).[4] While he emphasizes the use of page-references for clearly identifying pieces of empirical evidence, the underlying transparency principle ought to apply at least as much to the citation of works on which a scholar draws conceptually and theoretically.[5] And, in the quite different context of field research, Parkinson and Wood direct our attention to the ethical and practical importance of transparency toward one's research subjects—a form of scholarly openness that must begin to unfold long before the presentation of findings. We return to this concern below. For the moment, we wish to highlight our understanding of research transparency as encompassing issues that are logically prior to data-production, analysis, and data-sharing.

Production Transparency

Across the contributions there is relatively broad agreement on the importance of being transparent about the procedures used to collect or generate data. Bleich and Pekkanen, for instance, identify several features of interview-based research—including how interviewees were chosen, response rates and saturation levels achieved within interviewee categories, as well as interview formats and recording methods used—which researchers should report to make their results interpretable. Similarly, Romney, Stewart, and Tingley point to the importance of clearly defining the "universe" from which texts were chosen so that readers can appropriately assess inferences drawn from the analysis of those texts. In the context of ethnographic research, Cramer explains why detailed depictions of the context of her conversations with rural residents are cen-

[1] APSA 2012, 10 (section 6.2).

[2] APSA 2012, 10 (section 6.3).

[3] Lukes (1974), esp. 26ff, building on Gallie (1955-1956).

[4] Trachtenberg 2015, 13–14.

[5] Partly inspired by Trachtenberg's plea, we have asked all of our contributors to provide such pin-cites and we will require pin-cites for contributions to *QMMR* going forward.

tral to making sense of her claims. And Pachirat highlights interpretive ethnographers' rich accounts of the contours of their immersion, including attention to positionality and embodiment. Likewise, Davison discusses hermeneutical explanation as an enterprise in which the researcher brings to the surface for the reader her initial understandings and how those understandings shifted in the course of conversation with the material.

We also, however, observe important differences across the essays in how production transparency is conceptualized and operationalized. One divergence concerns the very category of "data production." For most positivist researchers,[6] it is quite natural to conceptualize a process of collecting evidence that unfolds prior to—or, at least, can be understood as separate from—the analysis of that evidence. From the perspective of interpretive work, however, a strict distinction between production and analytic transparency is less meaningful. To the extent that interpretivists reject the dualist-positivist ontological position that there is an empirical reality that is independent of the observer, "data" about the social and political world is not something that exists a priori and can be "collected" by the analyst.[7] In understanding observation and analysis as inextricably linked, the interpretive view also implies a *broader* understanding of transparency than that encompassed by the current DA-RT agenda. As Pachirat and Davison argue, interpretivism by its very nature involves forms of transparency—e.g., about the researcher's own prior understandings and social position, and how they may have influenced her engagement with the social world—that are not normally expected of positivist work.

Even where they are in broad agreement on principles, contributors give production transparency differing operational meanings, resulting from differences in the practical and ethical constraints under which different forms of research operate. All of the contributors who are writing from a positivist perspective, for instance, agree that information about sampling—i.e., selecting interview subjects, fieldwork locales, texts, or cases more generally—is important for assessing the representativeness of empirical observations and the potential for generalization. Yet, only for Bleich and Pekkannen (interview-based research), Romney, Stewart, and Tingley (computer-assisted textual analysis), and Wagemann and Schneider (QCA), is maximal explicitness about sampling/selection treated as an unconditional good. We see a quite different emphasis in those essays addressing field research outside stable, democratic settings. Parkinson and Wood (focused on ethnographic fieldwork in violent contexts) and Shih (focused on research in authoritarian regimes) point to important trade-offs between, on the one hand, providing a full account of how one's interviewees and informants were selected and, on the other hand, ethical obligations to research subjects and practical constraints arising from the need for access.[8] In some situations, ethical duties may place very stringent limits on transparency about sources, precluding even an abstract characterization of the individuals from whom information was received. As a colleague with extensive field research experience in non-democratic regimes pointed out to us: "If I attribute information, which only a limited number of people have, to a 'senior government official' in one of the countries where I work, their intelligence services can probably readily figure out my source, since they surely know with whom I've spoken."[9]

Also, as Cramer's discussion makes clear, forms of political analysis vary in the degree to which they aim to generate portable findings. For research traditions that seek to understand meanings embedded in particular contexts, the criteria governing transparency about case-selection will naturally be very different from the criteria that should operate for studies that seek to generalize from a sample to a broader population.

Our contributors also differ over how exactly they would like to see the details of data-production conveyed. Whereas many would like to see information about the evidence-generating process front-and-center, Trachtenberg argues that a finished write-up should generally *not* provide a blow-by-blow account of the reasoning and research processes that gave rise to the argument. A detailed account of how the empirical foundation was laid, he argues, would impede the author's ability to present the most tightly argued empirical case for her argument (and the reader's ability to benefit from it); it would also distract from the substantive issues being examined. An easy compromise, as we see it, would be the use of a methods appendix, much like the "Interview Methods Appendix" that Bleich and Pekkanen discuss in their contribution, which can provide detailed production-transparency information without disturbing the flow of the text.[10]

Qualitative Production Transparency for Quantitative Research

While our contributors have discussed matters of production transparency in the context of *qualitative* evidence, many of these same issues also logically arise for *quantitative* researchers using pre-existing datasets. The construction of a "quantitative" dataset always hinges on what are fundamentally qualitative judgments about measurement, categorization, and coding. Counting civilian casualties means deciding who counts as a "civilian";[11] counting civil wars means deciding not just what qualifies as "war" but also what qualifies as an internal war;[12] and measuring foreign direct investment means deciding what counts as "foreign" and "direct" as well as what

---

[6] Throughout this concluding essay, we use "positivism" or "positivist[ic]" broadly, to refer not just to pre-Popperian positivism but also the neopositivist tradition in the philosophy of science.

[7] See, e.g., Jackson 2008.

[8] We might add the consideration of the researcher's own safety, though none of our contributors emphasizes this point.

[9] Interview with Edmund Malesky, Duke University, 21 July 2015.

[10] For an example, see Büthe and Mattli 2011, 238–248.

[11] See, e.g., Lauterbach 2007; Finnemore 2013.

[12] See Sambanis 2004 vs. Fearon and Laitin 2003, esp. 76–78; Lawrence 2010.

counts as an "investment."[13] Such judgments involve the interpretation and operationalization of concepts as well as possibly the use of detailed case knowledge.[14] Qualitative judgments also enter into processes of adjustment, interpolation, and imputation typically involved in the creation of large datasets. And, naturally, the choices that enter into the construction of quantitative datasets—whether those choices are made by other scholars, by research assistants, or by employees of commercial data providers or government statistical offices—will be just as assumption- or value-laden, subjective, or error-prone as those made by qualitative researchers collecting and coding their own data.[15]

Current transparency guidelines for research using pre-existing datasets focus almost entirely on the disclosure of the source of the data (and on data access). But the information contained in a pre-existing dataset may not be rendered interpretable—that is, meaningful production transparency is not necessarily assured—simply by reference to the third party that collected, encoded/quantified, or compiled that information from primary or secondary texts, interviews, surveys, or other sources. Thus, in highlighting elements of the research process about which QMMR scholars need to be transparent, the contributors to this symposium also identify key questions that every scholar should ask—and seek to answer for her readers—about the generation of her data. And if the original source of a pre-existing dataset does not make the pertinent "qualitative" information available,[16] transparency in quantitative research should at a minimum entail disclosing that source's lack of transparency.

Analytic Transparency

The contributors to our symposium agree on the desirability of seeing researchers clearly map out the linkages between their observations of, or engagement with, the social world, on the one hand, and their findings and interpretations, on the other. Unsurprisingly, however, the operationalization of analytic transparency looks vastly different across analytical approaches.

The essays on QCA and on automated textual analysis itemize a number of discrete analytical decisions that, for each method, can be quite precisely identified in the abstract and in advance. Since all researchers employing QCA need to make a similar set of choices, Wagemann and Schneider are able to itemize with great specificity what researchers using this method need to report about the steps they took in moving from data to case-categorization to "truth tables" to conclusions. Romney, Stewart, and Tingley similarly specify with great precision the key analytical choices in computer-assisted text analysis about which transparency is needed. And Fairfield, in her disc-

ussion of Bayesianism, suggests several discrete inputs into the analysis (priors, likelihoods), which scholars who are using process-tracing can succinctly convey to the reader to achieve greater clarity about how they have drawn causal inferences from pieces of process-related evidence. In these approaches, it is not hard to imagine a reasonably standardized set of analytic-transparency requirements for each method that might be enforced by journal editors and reviewers.

Writing from an interpretive perspective, Pachirat endorses the general principle of analytic transparency in that he considers "leav[ing] up enough of the scaffolding in [a] finished ethnography to give a thick sense to the reader of how the building was constructed" as a key criterion of persuasive ethnography. Cramer likewise points to the importance of analytic transparency in her work, especially transparency about epistemological underpinnings. As an interpretive scholar traversing substantive terrain typically occupied by positivists—the study of public opinion—making her work clear and compelling requires her to lay out in unusual detail what precisely her knowledge-generating goals are and how she seeks to reach them. In contrast to authors working in positivist traditions, however, contributors writing on interpretive and hermeneutic approaches discuss general principles and logics of transparency, rather than precise rules. They do not seek to specify, and would likely reject, a "checklist" of analytical premises or choices that all scholars should disclose. Davison, in fact, argues that one of the defining characteristics of the hermeneutic approach is transparency regarding scholars' inherent inability to be fully aware of the intersubjective content of the conversation in which they are engaged.

Taking Analytic Transparency Further

Our authors also emphasize three aspects of analytic transparency that go beyond the current DA-RT agenda. First is transparency about *uncertainty*, a form of openness that does not frequently feature in any explicit way in qualitative research. In her discussion of process tracing, Fairfield examines the promise of Bayesian approaches, which are intrinsically probabilistic and thus provide a natural way to arrive at and express varying levels of certainty about propositions. Romney, Stewart, and Tingley, in their essay on automated textual analysis, highlight the importance of carrying over into our inferences any uncertainty in our measures. And Bleich and Pekkanen argue that scholars should disclose information about their interview-based research that allows the reader to arrive at an assessment of the confidence that she can have in the results. Researchers should routinely report, for instance, whether the interview process achieved "saturation" for a given category of actors and, when deriving a claim from interview responses, what proportion of the relevant interviewees agreed on the point. This kind of information is rarely provided in interview-based research at present, but it would be straightforward for scholars to do so.

If uncertainty in positivist work can be conceived of as a distribution or confidence interval around a point estimate, a more radical form of uncertainty—a rejection of fixed truth

---

[13] See Bellak 1998; Büthe and Milner 2009, 189f; Golub, Kauffmann, Yeres 2011; IMF, various years.

[14] See Adcock and Collier 2001.

[15] On this point, see also Herrera and Kapur 2007.

[16] See, for instance, the long-standing critiques of Freedom House's democracy ratings, such as Bollen and Paxton (2000); for an overview, see Büthe 2012, 48–50.

claims—is often central to the interpretive enterprise. As Davison writes of the hermeneutic approach: "There is no closing the gates of interpretation around even the most settled question."[17] The interpretivist's attention to researcher subjectivity and positionality implies a fundamental epistemological modesty that, as we see it, constitutes an important form of research openness.[18]

Second, transparency about the *sequence* of analytic steps is often lacking in both qualitative and quantitative research. In positivist studies, a concern with analytic sequence typically derives from an interest in distinguishing the empirical *testing* of claims from the *induction* of claims through exploratory analysis.[19] Scholars frequently default to the language of "testing" even though qualitative and multi-method research (and, we might add, quantitative research) frequently involves alternation between induction and deduction. In computer-assisted text analysis, for instance, processing procedures and dictionary terms are often adjusted after initial runs of a model. In QCA, conditions or cases might be added or dropped in light of early truth table analyses. As Romney, Stewart, and Tingley as well as Wagemann and Schneider emphasize, such a "back-and-forth between ideas and evidence"[20] is not merely permissible; it is often essential. Yet, they also argue, if readers are to be able to distinguish between empirically grounded hypotheses and empirically tested findings, then scholars must be more forthright about *when* their claims or analytical choices have been induced from the data and *when* they have been applied to or tested against fresh observations. The current push (largely among experimentalists) for pre-registration of studies and analysis plans[21] is an especially ambitious effort to generate such transparency by making it harder for researchers to disguise inductively generated insights as test results. Yet one can also readily imagine greater *explicitness* from researchers about analytic sequence, even in the absence of institutional devices like registration.[22]

[17] Davison 2015, 45.

[18] See also Pachirat's (2015) discussion of these issues.

[19] King, Keohane, and Verba 1994, 21–23, 46f. We focus here on positivist research traditions, but note that sequence is also important for non-positivists. Davison points to the importance, in hermeneutic work, of identifying the initial questions or perplexities that motivated the research, prior understandings revealed through conversation, and the new meanings received from this engagement. In this context, the sequence—the interpretive journey—in some sense *is* the finding.

[20] Ragin as quoted by Wagemann and Schneider 2015, 39.

[21] While this push has been particularly prominent among experimentalists (see Humphreys et al. 2013), it also has been advocated for other research approaches in which hypothesis-generation and data-analysis can be temporally separated. See, for instance, the 2014 call for proposals for a special issue of *Comparative Political Studies* (to be edited by Findley, Jensen, Malesky, and Pepinsky), for which review and publication commitments are to be made solely on the basis of research designs and pre-analysis plans. See http://www.ipdutexas.org/cps-transparency-special-issue.html.

[22] One distinct value of the time-stamped registration of a pre-analysis plan is that it makes claims about sequence more credible.

Third, the contributors to this symposium highlight (and differ on) the level of *granularity* at which the links between observations of the social world and conclusions must be spelled out. At the limit, analytic transparency with maximum granularity implies clarity about how each individual piece of evidence has shaped the researcher's claims. A maximally precise account of how findings were arrived at allows the reader to reason counterfactually about the results, asking how findings would differ if a given observation were excluded from (or included in) the analysis or if a given analytic premise were altered. In general, analytical approaches that rely more on mechanical algorithms for the aggregation and processing of observations tend to allow for—and their practitioners tend to seek—more granular analytic accounts. In QCA, for instance, while many observations are processed at once, the specificity of the algorithm makes it straightforward to determine, for a given set of cases and codings, how each case and coding has affected the conclusions. It would be far harder for an interpretive ethnographer to provide an account of her analytic process that made equally clear how her understandings of her subject matter turned on individual observations made in the field. Perhaps more importantly, as Cramer's and Pachirat's discussions suggest, such granularity—the disaggregation of the researcher's engagement with the world into individual pieces of evidence—would run directly against the relational and contextualized logics of interpretive and immersive forms of social research.

Moreover, Fairfield argues that, even where the scope for precision is high, there may be a tradeoff between granularity of transparency and fidelity to substance. In her discussion of process tracing, Fairfield considers varying degrees of formalism that scholars might adopt—ranging from clear narrative to formal Bayesianism—which make possible varying degrees of precision about how the analysis was conducted. If process tracing's logic can be most transparently expressed in formal Bayesian terms,[23] however, implementing maximally explicit procedures may force a form of false precision on researchers, as they seek to put numbers on informal beliefs. Further, Fairfield suggests, something may be lost in reducing an interpretable "forest" of evidence to the "trees" of individual observations and their likelihoods.[24]

Data Access

There is wide agreement on the virtues of undergirding findings and interpretations with rich empirical detail in written research outputs.[25] However, we see great divergence of views among symposium contributors concerning the meaning and operationalization of data access. It is perhaps unsurprising

[23] See, e.g., Bennett 2015; Humphreys and Jacobs forthcoming.

[24] See also Büthe's (2002, 488f) discussion of the problem of reducing the complex sequence of events that make up a macro-historical phenomenon to a series of separable observations.

[25] The essays by Cramer, Fairfield, Pachirat, and Trachtenberg, for instance, discuss (variously) the importance of providing careful descriptions of case or fieldwork context and extensive quotations from observed dialogue, interviews, or documents.

that data access might be a particular focus of controversy; it is, in a sense, the most demanding feature of the DA-RT initiative. Whereas production and analytic transparency admit of widely varying interpretations, the principle of data access asks all empirical researchers to do something quite specific: to provide access to a full and unprocessed record of "raw" empirical observations, such as original source documents, full interview transcripts, or field notes.

Several of the symposium contributors—including Bleich and Pekkanen (interviews), Romney, Stuart, and Tingley (computer-assisted text analysis), Wagemann and Schneider (QCA), and Trachtenberg (documentary evidence)—call for researchers to archive and link to, or otherwise make available, as much of their raw data as feasible. These scholars have good reasons for wanting to see this level of data access: most importantly, that it enhances the credibility and interpretability of conclusions. And where data access involves access to the full empirical record (as in some examples provided by Romney, Stuart, and Tingley or when it entails access to complete online archives), it can help identify and correct the common problem of confirmation bias[26] by allowing readers to inspect pieces of evidence that may not have featured in the researcher's account or analysis. In addition, data access promises to reduce the frequency with which suspicions of misbehavior or questionable practices cannot be dispelled because the referenced sources or data can be neither found nor reproduced. Broad data access could also facilitate the kind of scholarly adversarialism called for by Trachtenberg, where researchers test one another's arguments at least as frequently as their own.[27] In sum, although they acknowledge that legal restrictions and ethical obligations may sometimes make full data disclosure impossible or unwise, these authors express a default preference for maximally feasible data-sharing.

Several other contributors, however, raise two important and distinct objections—one intellectual and one ethical—to data access as the presumptive disciplinary norm. The critique of data access on intellectual grounds—prominent in Pachirat's and Cramer's contributions—takes issue with the very concept of "raw data." These authors argue, on the one hand, that from an interpretive-ethnographic standpoint, there is no such thing as pre-analytic observations. Field notes, for example, are *more* than "raw data" because, as Pachirat argues, they already reflect both "the intersubjective relations and the implicit and explicit interpretations" that will inform the finished work.[28] At the same time, ethnographic transcripts and field

notes are *less* than "raw data" because they are torn from context. Cramer points to the many features of conversations and contexts that are not recorded in her transcripts or notes, but are critical to the meaning of what is recorded. As she puts it, her raw data do not consist of "my transcripts and field-notes....The raw data exists in the act of spending time with and listening to people."[29]

The second, ethical critique of data access figures most prominently in the essays by Parkinson and Wood on intensive field research in contexts of political violence, Shih on research in non-democratic regimes, and Cramer on the ethnographic, interpretive study of public opinion. In such contexts, these authors argue, ethical problems with full data access are not the exception; they are the norm. And, importantly, these are not challenges that can always be overcome by "informed consent." For Cramer, the idea of posting full transcripts of her conversations with rural residents would fundamentally undermine the ethical basis of those interactions. "I am able to do the kind of work I do," Cramer writes, "because I am willing to put myself out there and connect with people on a human level…. If the people I studied knew that in fact they were not just connecting with me, but with thousands of anonymous others, I would feel like a phony, and frankly would not be able to justify doing this work."[30] The type of immersive inquiry in which Cramer engages is, in a critical sense, predicated on respect for the *privacy* of the social situation on which she is intruding. Even where subjects might agree to full data-sharing, moreover, *informed* consent may be an illusory notion in some research situations. As Parkinson and Wood point out, a pervasive feature of research in contexts of political violence is the difficulty of fully forecasting the risks that subjects might face if their words or identities were made public.[31] We discuss the implications of these ethical challenges further below.

Replication, Reproducibility, and the Goals of Research Transparency

As noted in our introduction to the symposium, calls for greater research transparency in Political Science have been motivated by a wide range of goals. Lupia and Elman in their essay in the January 2014 DA-RT symposium in *PS*—like the contributors to our symposium—elaborate a diverse list of potential benefits of transparency. These include facilitating diverse forms of evaluation of scholarly claims, easing the sharing of knowledge across research communities unfamiliar with one another's methods or assumptions, enhancing the usefulness of research for teaching and for non-academic audiences, and making data available for secondary analysis.[32]

Enabling replication, as a particular form of evaluation, is among the many declared objectives of the push for research openness.[33] Replication has attracted considerable controversy

---

[26] Confirmation bias refers to the tendency to disproportionately search for information that confirms one's prior beliefs and to interpret information in ways favorable to those beliefs. It has long been emphasized by political psychologists as a trait of decisionmakers (see, e.g., Jervis 1976, esp. 128ff, 143ff, 382ff; Tetlock 2005) but is equally applicable to researchers.

[27] Trachtenberg 2015, 16. For a similar argument about the benefits of science as a collective and somewhat adversarial undertaking, from a very different tradition, see Fiorina 1995, 92. Note that vibrant adversarialism of this kind requires journals to be willing to publish null results.

[28] Pachirat 2015, 30.

[29] Cramer 2015, 20.

[30] Cramer 2015, 19.

[31] Parkinson and Wood 2015, 23–24, 26.

[32] Lupia and Elman 2014, 20.

[33] Interestingly, King (1995: esp. 444f) motivated his early call for more replication in quantitative political science by reference to nearly

and attention, not least since the 2012 revision of the APSA *Ethics Guide* includes an obligation to enable one's work to be "tested or replicated."[34] Arguments made by our contributors have interesting implications for how we might conceive of the relationship between transparency and forms of research evaluation.

Replication so far has not featured nearly as prominently in discussions of transparency for qualitative as for quantitative research.[35] Among the articles in this symposium, only Romney, Stuart, and Tingley explicitly advocate replicability as a standard by which to judge the transparency of scholarship. Notably, their essay addresses a research tradition—the computer-assisted analysis of texts—in which the literal reproduction of results is often readily achievable and appropriate, given the nature of the data and the analytic methods. (The same could perhaps be said of QCA.) At the other end of the spectrum, authors writing on immersive forms of research explicitly reject replication as a relevant goal,[36] and for research like Davison's (hermeneutics), it is clearly not an appropriate evaluative standard.[37]

Many of the essays in our symposium, however, seem to endorse a standard that we might call enabling "replication-in-thought": the provision of sufficient information to allow readers to trace the reasoning and analytic steps leading from observation to conclusions, and think through the processes of observation or engagement. Replication-in-thought involves the reader asking questions such as: Could I in principle imagine employing the same procedures and getting the same results? If I looked at the evidence as presented by the author, could I reason my way to the same conclusions? Replication-in-thought also allows a reader to assess how the researcher's choices or starting assumptions might have shaped her conclusions.

This approach seems to undergird, at least implicitly, much of the discussions by Fairfield, Parkinson and Wood, Shih, and Trachtenberg; it is central to what Bleich and Pekkanen call "credibility." This kind of engagement by the audience, moreover, is not unique to positivist scholarship. Pachirat writes of the importance of interpretive ethnographies providing enough information about how the research was conducted and enough empirical detail so that the reader can interrogate and challenge the researcher's interpretations. The transparency and data demands for meaningful replication-in-thought might also be somewhat lower than what is required for literal replication.

We also think that there may be great value—at least for positivist qualitative scholars—in putting more emphasis on a related evaluative concept that is common in the natural sciences: the *reproducibility of findings*.[38] Replication, especially in the context of quantitative political science, is often understood as generating the same results by issuing the same commands and using the same data as did the original researcher. Replication in political science sometimes also extends to the application of new analytical procedures, such as robustness tests, to the original data. Replication in the natural sciences, by contrast, virtually always means re-generating the data as well as re-doing the analysis. This more ambitious approach tests for problems not just in the execution of commands or analytic choices but also in the data-generation process.

Of course, re-generating an author's "raw data" with precision may be impossible in the context of much social research, especially for small-*n* qualitative approaches that rely on intensive fieldwork and interactions with human subjects. Yet the problem of precisely replicating a study protocol is not limited to the social sciences. Much medical research, for instance, is also highly susceptible to a broad range of contextual factors, which can make exact replication of the original data terribly difficult.[39]

Some researchers in the natural sciences have thus turned from the replication of study protocols to the alternative concept of the reproducibility of findings.[40] Reproducibility entails focusing on the question: Do we see an overall consistency of results when the same research question is examined using different analytical methods, a sample of research subjects recruited elsewhere or under somewhat different conditions, or even using an altogether different research design?

A focus on reproducibility—which begins with gathering afresh the empirical information under comparable but not identical conditions—may solve several problems at once. It mitigates problems of data access in that sensitive information about research subjects or informants in the original study need not be publicly disclosed, or even shared at all, to achieve an empirically grounded re-assessment of the original results. As Parkinson and Wood put it with respect to fieldwork, scholars "expect that the over-arching findings derived from good fieldwork in similar settings on the same topic should converge significantly."[41] A reproducibility standard can also cope with the problem that precise repetition of procedures and observations will often be impossible for qualitative research.

---

the same set of benefits of research transparency noted in recent work.

[34] APSA 2012, 9.

[35] See, e.g., Elman and Kapiszewski 2014. Moravcsik (e.g., 2010) is rather unusual in promoting partial qualitative data access through hyperlinking (see also Trachtenberg 2015) explicitly as a means to encourage replication analyses in qualitative research.

[36] See, in particular, Pachirat, Parkinson and Wood, and Cramer.

[37] Lupia and Elman similarly note that members of research communities that do not assume inferential procedures to be readily repeatable "do not [and should not be expected to] validate one another's claims by repeating the analyses that produced them" (2014, 22).

[38] Saey 2015, 23. Unfortunately, the terms "replication" and "reproducibility" do not have universally shared meanings and are sometimes used inconsistently. We have therefore tried to define each term clearly (below) in a manner that is common in the social sciences and consistent with the predominant usage in the natural sciences, as well as with the definition advocated by Saey.

[39] Djulbegovic and Hozo 2014. See also Bissell 2013.

[40] Saey 2015, 23.

[41] Parkinson and Wood also draw out this point. Similar arguments for the evaluation of prior findings via new data-collection and analyses rather than through narrow, analytic replication are made by Sniderman (1995, 464) and Carsey (2014, 73) for quantitative research.

A focus on reproducing findings in multiple contexts and with somewhat differing designs can, further, help us learn about the contextual conditions on which the original findings might have depended, thus uncovering scope conditions or causal interactions of which we were not previously aware.[42] Further, if we conceive of the characteristics of the researcher as one of the pertinent contextual conditions, then shifting the focus to reproducibility might even allow for common ground between positivist researchers, on the one hand, and on the other hand interpretivist scholars attentive to researcher positionality and subjectivity: When different researchers arrive at different results or interpretations, as Parkinson and Wood point out, this divergence "should not necessarily be dismissed as a 'problem,' but should be evaluated instead as potentially raising important questions to be theorized."[43]

Broader Implications: Achieving Research Transparency and Maximizing Its Benefits

The issues raised in this symposium, and in the larger literature on research transparency and research ethics, have a number of implications for the drive for openness in Political Science. In this section, we draw out some of these implications by considering several ways in which scholarly integrity, intellectual pluralism, and research ethics might be jointly advanced in our discipline. We discuss here four directions in which we think the transparency agenda might usefully develop: toward the elaboration of differentiated transparency concepts and standards; a more explicit prioritization of human-subject protections; a realignment of publication incentives; and more robust training in research ethics. We also draw attention to transparency's considerable costs to scholars and their programs of research—a price that may well be worth paying, but that nonetheless merits careful consideration.

Differentiated Transparency Concepts and Standards

When the APSA's Ad Hoc Committee on Data Access and Research Transparency and its subcommittees drew up their August 2013 Draft Guidelines, they differentiated between a quantitative and a qualitative "tradition."[44] The elaborate and thoughtful Guidelines for Qualitative Research and the accompanying article about DA-RT "in the Qualitative Tradition"[45] went to great lengths to be inclusive by largely operating at a high level of abstraction. Elman and Kapiszewski further recognized the need to distinguish both between broad approaches (at a minimum between positivist and non-positivist inquiry) and between the plethora of specific non-statistical methods that had been subsumed under the "qualitative" label. The actual October 2014 "DA-RT Statement" similarly acknowledges diversity when it notes that "data and analysis take diverse forms in different traditions of social inquiry" and that therefore "data access and research transparency pose different challenges for [the many different research] traditions" and that "the means for satisfying the obligations will vary correspondingly."[46] The DA-RT Statement, however, does not itself work out this differentiation. The editorial commitments in the Statement are formulated in quite general terms that are "intended to be inclusive of specific instantiations."

The discussion that has unfolded in these pages underscores the critical importance of following through on the promise of differentiation: of developing transparency concepts and standards that are appropriate to the particular research traditions encompassed by the discipline. Computer-assisted content analysis, for instance, requires a series of very particular analytical decisions that can be highly consequential for what a scholar may find and the inferences she might draw. Understanding, assessment, and replication alike require scholars using such methods to provide all of the particular pieces of information specified by Romney, Stewart and Tingley in their contribution to the symposium. Yet, requests for that same information would make no sense for an article based on qualitative interviews where the analysis does not involve a systematic coding and counting of words and phrases but a contextualized interpretation of the meaning and implications of what interview subjects said. If journals aim to implement the DA-RT principles in a way that can truly be applied broadly to forms of political inquiry, they will need to develop or adopt a set of differentiated "applied" standards for research transparency.

One danger of retaining uniform, general language is that it will simply remain unclear—to authors, reviewers, readers—what kind of openness is expected of researchers in different traditions. What information about production or analysis are those doing interviews, archival research, or ethnography expected to provide? To achieve high transparency standards, moreover, researchers will need to know these approach-specific rules well in advance of submitting their work for publication—often, before beginning data-collection. A second risk of seemingly universal guidelines for contributors is the potential privileging of research approaches that readily lend themselves to making all cited evidence available in digital form (and to precisely specifying analytic procedures) over those approaches that do not. We do not think that DA-RT was intended to tilt the scholarly playing field in favor of some scholarly traditions at the expense of others. But uniform language applied to all empirical work may tend to do just that.

Perhaps most significantly, journals may need to articulate a separate set of transparency concepts and expectations for positivist and interpretivist research traditions. As the contributions by Cramer, Davison, and Pachirat make clear, scholars in the interpretivist tradition are strongly committed to research transparency but do not subscribe to certain epistemological assumptions underlying DA-RT. The differences between positivist and interpretivist conceptions of research transparency thus are not just about the details of implementation: they get to fundamental questions of principle, such as

---

[42] Parkinson and Wood 2015, 25. See also Saey 2015, 23.

[43] Parkinson and Wood 2015, 25.

[44] The Draft Guidelines were published as appendices A and B to Lupia and Elman 2014.

[45] Elman and Kapiszewski 2014.

[46] DA-RT 2014, 1.

whether data access should be required. Without taking this diversity into account, undifferentiated standards might de facto sharply reduce the publication prospects of some forms of interpretive scholarship at journals applying DA-RT principles.[47]

We find it hard to see how a single, complete set of openness principles—in particular, when it comes to data access—could apply across positivist and interpretivist traditions in a manner that respects the knowledge-generating goals and epistemological premises underlying each. We thus think it would be sensible for publication outlets to develop distinct, if partially overlapping, transparency standards for these two broad forms of scholarships. Most simply, authors could be asked to self-declare their knowledge-generating goals at the time of submission to a journal and in the text of their work: Do they seek to draw observer-independent descriptive or causal inferences? If so, then the data-generating process is indeed an extractive one, and—within ethical, legal, and practical constraints—access to the full set of unprocessed evidence should be expected. By contrast, if scholars seek interpretations and understandings that they view as tied to the particularities of their subjective experience in a given research context, then transparency about the constitutive processes of empirical engagement and interpretation would seem far more meaningful than "data access." At the same time, even ethnographic scholars might still be expected to make available evidence of certain "brute facts" about their research setting that play a role in their findings or interpretations. As compared to a uniform set of requirements, such an arrangement would, we think, match the scholar's obligations to be transparent about her research process more closely to the knowledge claims that she seeks to make on the basis of that process.

The Primacy of Research Ethics

Parkinson and Wood as well as Shih emphasize the dire risks that local collaborators, informants, and research subjects of scholars working in non-democratic or violent contexts might face if field notes or interview transcripts were made public. These risks, Parkinson and Wood argue further, are often very difficult to assess in advance, rendering "informed" consent problematic. We concur. In such situations, researchers' ethical obligations to protect their subjects will necessarily clash with certain principles of transparency. And it should be uncontroversial that, when in conflict, the latter must give way to the former.

In the 2012 Revision of the APSA *Ethics Guide*, which elevated data access to an ethical obligation, human subjects protections clearly retained priority. But as Parkinson and Wood point out, the October 2014 DA-RT Statement does not appear to place ethical above transparency obligations, as scholars who withhold data in order to protect their research subjects

must ask for editorial exemption before having their papers reviewed. In its framing, the DA-RT language appears to place at a presumptive disadvantage any research constrained from full disclosure by ethical or legal obligations to human subjects. The DA-RT formulation also places a difficult moral judgment in the hands of a less-informed party (the editor) rather than a more informed one (the researcher). And it creates a worrying tension between the professional interests of the researcher and the interests of her subjects.

We want to emphasize that we do not believe that DA-RT advocates intend to see data access trump the ethical treatment of human subjects, nor that they expected the October 2014 DA-RT Statement to have such an effect. As Colin Elman put it in an interview with us: "I would be truly astonished if, as a result of DA-RT, journals that now publish qualitative and multi-methods research begin to refuse to review manuscripts because they employ data that are genuinely under constraint due to human subjects concerns."[48] At the same time, we see concerns such as those articulated by Parkinson and Wood as reasonable, given that the DA-RT language to which journals have committed themselves does not explicitly defer (nor even make reference) to the APSA *Ethics Guide's* prioritization of human subject protection, and that most of the signatory journals have no formal relationship to the APSA or to any other structure of governance that could be relied upon to resolve a conflict of norms.

We believe that a few, simple changes in the wording of the DA-RT Statement might go a substantial way toward mitigating the concerns of scholars conducting research in high-risk contexts. The DA-RT Statement could be amended to state clearly that ethical and legal constraints on data access take priority and that editors have an obligation to respect these constraints without prejudice to the work being considered. In other words, editors should be expected to require maximal data access *conditional on the intrinsic shareability* of the data involved. The goal here should be maximizing the useful sharing of data from research that merits publication—rather than favoring for publication forms of social research with intrinsically more shareable data.

Of course, one clear consequence of exempting some forms of scholarship from certain transparency requirements would be to elevate the role of trust in the assessment and interpretation of such research. It would require readers to defer somewhat more to authors in assessing whether, e.g., an authoritarian regime might be manipulating the information to which the researcher has access; whether the sources of her information are disinterested and reliable; whether self-censorship or concern for local collaborators and informants is seriously affecting the questions asked and hypotheses examined; and whether the evidence presented is representative of the full set of observations gathered.

At the same time, a focus on *reproducible findings*—as we have advocated above—may soften this tradeoff. Full data access and production transparency are critically important if

---

[47] And even if a journal sees itself as an outlet for positivist work only, it is surely better that it be at least, well, transparent about it, so that expectations can be suitably adjusted and so that scholars, especially junior scholars, do not waste their time on submissions that stand no chance of success.

[48] Phone interview with Colin Elman, Syracuse University, 6 July 2015.

evaluation rests on the re-analysis of existing data. It matters considerably less if researchers are willing to probe prior findings through new empirical studies carried out in the same or comparable settings.[49] Fresh empirical work is, of course, more costly than analytic replication with old data. But it surely offers a more robust form of evaluation and, as argued above, broader opportunities for learning.

Realigning Publication Incentives to Achieve the Full Benefits of Openness

To date, DA-RT proponents have asked journal editors to help advance the cause of research openness by creating a set of transparency and data access requirements for submitted or published manuscripts. We think there are at least two additional roles—going beyond the use of requirements or negative sanctions—that journals can play in advancing the broader aim of research integrity, especially in the context of positivist research.[50]

First, journals could do more to reward efforts to directly evaluate the findings of previous studies. Research transparency is not an end in itself. Rather, it is a means to (among other things) the evaluation of research. At the moment, however, outside the graduate methods classroom, replication, reproduction, robustness tests, and other evaluative efforts are not in themselves highly valued in our discipline. Journals rarely publish the results of such evaluative efforts—part of a general tendency (also noted by Trachtenberg) to discount work that "merely" tests an existing theory or argument without making its own novel theoretical contribution.[51] We believe journals should complement requirements for greater transparency with a commitment to publish well-executed efforts to replicate or otherwise evaluate significant, prior results, especially high-quality efforts to reproduce prior findings with new data (including when those efforts succeed).

Relatedly, journals (and other publication outlets) could advance the cause of research integrity by taking more seriously findings that call previously published research into question. Note the contrast in this regard between recent developments in the life sciences and current practice in Political Science. Medical and related scientific journals are increasingly asserting for themselves the right to issue "statements of concern" and even to retract published pieces—without the consent of the authors, if necessary—if a thorough investigation reveals grave errors, highly consequential questionable research practices, or scientific misconduct.[52] By contrast, we are not aware of any cases where even dramatic failures of straight replication have resulted in retractions of, or editorial statements-of-concern about, published articles in Political Science.[53] The risk of drawn-out legal disputes has been one important reason why journals (and the publishers and professional associations that stand behind them) have shied away from calling out discredited findings.[54] Political Science journals might therefore want to follow the lead of medical journals, many of which ask authors to explicitly grant editors the right to issue statements of concern or to retract an article when serious problems are uncovered (and to do so without the authors' consent).

Second, journals can advance the cause of transparency by reorienting some currently misaligned incentives. Specifically, journals could directly reduce the temptation to disguise exploratory research as "testing" by demonstrating greater willingness to publish openly inductive work—a form of scholarship that is widely understood to play a critical role in advancing research agendas.[55] Similarly, editors could mitigate pressures to selectively report evidence or cherry-pick model specifications through greater openness to publishing mixed and null results of serious tests of plausible theories.

In sum, many of the unwelcome scholarly tendencies that transparency rules seek to restrain are a product of our own disciplinary institutions. Transparency advocates will thus be swimming against strong currents until we better align professional rewards with our discipline's basic knowledge-generating goals.

Ethics Training and Professional Norms

Beyond rules and incentives, we might also look to professional norms. In ongoing research, Trisha Phillips and Franchesca Nestor find that training in research ethics is much less commonly included among formal program requirements or mentioned in course descriptions for Political Science graduate programs than in other social sciences.[56] As we have discussed, certain operational tensions can arise between a researcher's ethical (and legal) obligations to human subjects and transparency requirements. At the level of professional *training*, however, the two goals could be mutually reinforcing and jointly conducive to scholarly integrity.

First, better training in research ethics could advance the cause of transparency by allowing political researchers to make

---

[49] Note that this is common practice in, for instance, the life sciences, where any "replication" is understood to mean repeating all stages of the research, including generating the data. As a consequence, human subject protections are generally not considered a significant impediment, even for "replication," even though the human subjects protections required by medical IRBs are often much stricter than for non-medical research, as research subjects in the life science might be exposed to serious material, psychological, or physical risks if identified.

[50] As Baldwin (1989 (1971)) pointed out, positive and negative sanctions have different normative appeal and differ in how they work and how well they are likely to work as ways to exert influence.

[51] This tendency is arguably weaker in the world of quantitative causal inference, where novel research designs for testing existing claims are sometimes valued in themselves, than in qualitative work.

[52] Resnik, Wager, and Kissling 2015.

[53] For quantitative work, the inability or failure to replicate is reported all too frequently, suggesting strongly that the lack of notices of concern and the lack of retractions in Political Science is not indicative of the absence of a problem. As we note in our Introduction, qualitative work might be less susceptible to some problems, but overall is surely not free of important errors.

[54] Wager 2015.

[55] See also Fenno 1986; Rogowski 1995.

[56] Phone interview with Trisha Phillips, West Virginia University, 16 July 2015, and Phillips and Nestor 2015.

more informed and confident decisions about ethical considerations. As Shih and Parkinson and Wood note, while human subject protections and related ethical obligations must be taken very seriously, we need not categorically consider all aspects of our field research confidential and all of our interviews non-attributable. Better training in research ethics should allow political scientists to avoid overly sweeping non-disclosure commitments simply as a default position that eases IRB approval. Nor should we see IRB requirements as the maximum required of us ethically. Scholars-in-training need guidance and a conceptual apparatus for thinking through (and explaining to editors!) the ethical dilemmas that they face.

Second, better training in research ethics, broadly conceived, could readily complement transparency initiatives in their efforts to mitigate research misconduct and enhance the value of scholarly findings. Transparency rules are designed to reduce the incentives for scholars to engage in fabrication, falsification, plagiarism, and "questionable research practices"[57] such as the selective reporting of supportive empirical findings or the biased interpretation of sources. They do so by increasing the likelihood that such actions will be disclosed. Yet, we are more likely as a profession to achieve the ultimate ends of broadly reliable research findings if we underwrite rules and sanctions with *norms* of research integrity. We presume that nearly all graduate students know that making up your data is plain wrong. But we wonder whether they are currently encouraged to think carefully about what counts as the unethical manipulation of results. To be sure, neither transparency nor training in research ethics will safeguard against intentional misconduct by committed fraudsters.[58] But, particularly in an environment in which Ph.D. students are counseled to aggressively pursue publication, more robust ethics training could help dispel any notion that it is acceptable to selectively present or omit relevant empirical information in the service of dramatic findings. Meanwhile, greater transparency could produce a growing body of published research that demonstrates such behavior *not* to be the norm.

Costs and Power in the Drive for Greater Transparency

Beyond the important practical and ethical issues highlighted in the contributions to our symposium, demands for research transparency also need to consider issues of cost and power. As research on technology governance shows, raising standards can have protectionist effects, creating barriers to entry that reinforce existing power- and market-structures, protecting the established and well-resourced to the detriment of new and under-resourced actors.[59] A number of our PhD students and untenured colleagues have raised similar concerns about the drive for greater research transparency: While document-

ing every step in the process of designing and executing a research project—in a format that everyone can access and understand—surely will lead to better research, ever-increasing demands for such documentation also imposes upon PhD students and junior faculty a set of burdens not faced by senior scholars when they were at the early stages of their careers. These are of course also the stages at which scholars are typically most resource-strapped and facing the most intense time-pressure to advance projects to publication. We submit that such concerns deserve to be taken very seriously (even if any "protectionist" effects are wholly unintended).

Qualitative scholars need to be especially attentive to this issue, because of the kinds of analytic reasoning and data that their work typically entails. One challenge arises from the diversity of types of evidence and, hence, the multiple distinct lines of inferential reasoning upon which a single case study often relies. In our symposium, Fairfield discusses the challenge in the context of trying to meticulously document the process-tracing in her work. As she recounts, outlining the analytic steps for a *single* case study in explicit, Bayesian terms required 10,000 words; doing so for all cases reported in her original *World Development* article[60] would have required nearly a book-length treatment.[61] As valuable as such explicitness may be, publishing an article based on case study research should not require a supporting manuscript several times the length of the article itself.

Qualitative data access can be similarly costly. Consider Snyder's recent effort to retrofit chapter 7 of his *Ideology of the Offensive*[62] with hyperlinks to copies of the source documents from the Soviet-Russian archives that had provided much of the empirical support regarding the "cult of the offensive" in the Russia case.[63] His notes from his visit more than 30 years ago to the Moscow archives—none of which allowed taking copies of materials with him at the time—were so good (and preserved!) that he was able to provide a virtually complete set of the cited source documents. But as he himself points out, this was "a major enterprise" that involved hiring a researcher in Moscow and a graduate research assistant for several months.[64]

Some of the costs of transparency can be greatly reduced by anticipating the need for greater openness from the start of a research project. This makes it all the more important to expose our PhD students to, and involve them in, debates over changing scholarly norms and practices early on. Technological innovations—e.g., software such as NVivo, which greatly facilitates documenting the analysis of texts, and the free tools developed for the new Qualitative Data Repository[65]—can also reduce transparency costs, and they might even increase the efficiency of some forms of qualitative inquiry.

---

[57] Smith 2008.

[58] Due to system effects (Jervis 1997), transparency might in fact lead to more elaborate frauds.

[59] See, e.g., Besen and Farrell 1994; Büthe and Mattli 2011, esp. 220–226; Henson and Loader 2001; Katz and Shapiro 1985; Maskus, Otsuki and Wilson 2005; and Rege, Gujadhur, and Franz 2003.

[60] Fairfield 2013.

[61] Fairfield 2015, 50; Fairfield and Charman 2015. See also Saunders (2014) thoughtful discussion of the issue of costs.

[62] Snyder 1984.

[63] See Snyder 2015.

[64] Snyder 2014, 714.

[65] See https://qdr.syr.edu (last accessed 7/20/2015).

Yet, the time and expense required to meet the new transparency demands will inevitably remain very high for many forms of research. The implementation of more ambitious openness standards will thus require careful consideration of transparency's price. Put differently, research transparency is a high-value but costly intellectual good in a world of scarcity. Scholars' transparency strategies should, like all research-design choices, be weighed in light of real resource constraints. Researchers should not be expected to achieve perfect transparency where doing so would impose prohibitive burdens on worthwhile research undertakings, stifle the dissemination of significant findings, or yield modest additional gains in interpretability. Rather, as with other matters of research design, scholars should be expected to do all that they reasonably can to establish the bases of their claims and to address potential objections, doubts, and ambiguities. We agree with DA-RT proponents that most of us can reasonably do a lot more to be transparent than we have been doing so far, and that standards have a critical role to play in generating greater openness. We will need to be careful, however, to apply transparency standards in ways that promote the ultimate goal of Political Science scholarship: advancing our understanding of substantively important issues.[66]

An Ongoing Conversation

The contributors to this symposium have sought to advance our understanding of research transparency in qualitative and multi-method research. They have done so by discussing the goals of transparency, exploring possibilities for achieving greater transparency, and probing the associated risks and limitations for a broad range of specific research traditions in Political Science.

Our contributors have approached the transparency agenda in a constructive spirit, even when voicing strong concerns and objections. Several of the articles have put forth detailed, specific proposals for how to achieve greater transparency at reasonable cost and with minimal drawbacks. Some have elaborated understandings of transparency in the context of their research traditions and epistemological and ontological commitments. Other authors have identified particular research contexts in which research transparency entails especially steep trade-offs and must be balanced against other, ethical and practical considerations. And in this concluding essay, we have sought to sketch out a few proposals that might address some of the issues identified by our contributors and more broadly advance the goals of scholarly integrity.

The symposium's goal has been to open and invite a broader conversation about research transparency in Political Science, rather than to close such a conversation by providing definitive answers. On one level, this is a conversation marked by starkly divergent perspectives on the nature of social inquiry, framed by uneasy disciplinary politics. On another level,

however, we see wide agreement on important "meta-standards" and on the core scholarly and ethical values at stake. And while trade-offs abound, we see no fundamental conflict between the pursuit of more transparent political analysis and the defense of intellectual pluralism in the discipline. We are thus cautiously optimistic about the prospects for progress through collective deliberation on the challenges highlighted by this collection of essays. We encourage members of the QMMR community, whatever their view of the issues discussed here, to fully engage in that conversation.

References

Adcock, Robert, and David Collier. 2001. "Measurement Validity: A Shared Standard for Qualitative and Quantitative Research." *American Political Science Review* vol. 95, no. 3: 529–546.

APSA, Committee on Professional Ethics, Rights and Freedoms. 2012. *A Guide to Professional Ethics in Political Science*. Second Edition. Washington, D.C.: APSA. (Available at http://www.apsanet.org/portals/54/Files/Publications/APSAEthicsGuide2012.pdf, last accessed July 16, 2015.)

Baldwin, David A. 1989 (1971). "The Power of Positive Sanctions." (First published *Journal of Conflict Resolution* vol. 15, no. 2 (1971): 145–155.) In *Paradoxes of Power*. New York: Basil Blackwell: 58–81.

Bellak, Christian. 1998. "The Measurement of Foreign Direct Investment: A Critical Review." *International Trade Journal* vol. 12, no. 2: 227–257.

Bennett, Andrew. 2015. "Disciplining Our Conjectures: Systematizing Process Tracing with Bayesian Analysis." In *Process Tracing in the Social Sciences: From Metaphor to Analytic Tool*, edited by Andrew Bennett and Jeffrey Checkel. New York: Cambridge University Press: 276–298.

Besen, Stanley, and Joseph Farrell. 1994. "Choosing How to Compete: Strategies and Tactics in Standardization." *Journal of Economic Perspectives* vol. 8, no. 2 (Spring 1994): 117–131.

Bissell, Mina. 2013. "Comment: The Risks of the Replication Drive." *Nature* no. 503 (21 November 2013): 333–334.

Bleich, Erik, and Robert J. Pekkanen. 2015. "Data Access, Research Transparency, and Interviews: The Interview Methods Appendix." *Qualitative and Multi-Method Research: Newsletter of the American Political Science Association's QMMR Section* vol. 13, no. 1: 8–13.

Bollen Kenneth A. and Pamela Paxton. 2000. "Subjective Measures of Liberal Democracy." *Comparative Political Studies* vol. 33, no. 2: 58–86.

Büthe, Tim. 2002. "Taking Temporality Seriously: Modeling History and the Use of Narratives as Evidence." *American Political Science Review* vol. 96, no. 3: 481–494.

———. 2012. "Beyond Supply and Demand: A Political-Economic Conceptual Model." In *Governance by Indicators: Global Power through Quantification and Rankings*, edited by Kevin Davis, et al. New York: Oxford University Press: 29–51.

Büthe, Tim, and Walter Mattli. 2011. *The New Global Rulers: The Privatization of Regulation in the World Economy*. Princeton: Princeton University Press.

Büthe, Tim, and Helen V. Milner. 2009. "Bilateral Investment Treaties and Foreign Direct Investment: A Political Analysis." In *The Effect of Treaties on Foreign Direct Investment: Bilateral Investment Treaties, Double Taxation Treaties, and Investment Flows*, edited by Karl P. Sauvant and Lisa E. Sachs. New York: Oxford University Press: 171–225.

---

[66] On the importance of maintaining a focus on the substantive contribution of social science research, see, e.g., Isaac (2015, 270f, 297f); Nye 2009; Pierson and Skocpol (2002, esp. 696–698); and Putnam (2003).

Carsey, Thomas M. 2014. "Making DA-RT a Reality." *PS: Political Science and Politics* vol. 47, no. 1 (January 2014): 72–77.

Cramer, Katherine. 2015. "Transparent Explanations, Yes. Public Transcripts and Fieldnotes, No: Ethnographic Research on Public Opinion." *Qualitative and Multi-Method Research: Newsletter of the American Political Science Association's QMMR Section* vol. 13, no. 1: 17–20.

"Data Access and Research Transparency (DA-RT): A Joint Statement by Political Science Journal Editors." 2014. At http://media. wix.com/ugd/fa8393_da017d3fed824cf587932534c860ea25. pdf (last accessed 7/10/2015).

Davison, Andrew. 2015. "Hermeneutics and the Question of Transparency." *Qualitative and Multi-Method Research: Newsletter of the American Political Science Association's QMMR Section* vol. 13, no. 1: 43–46.

Djulbegovic, Benjamin, and Iztok Hazo. 2014. "Effect of Initial Conditions on Reproducibility of Scientific Research." *Acta Informatica Medica* vol. 22, no. 3 (June 2014): 156–159.

Elman, Colin, and Diana Kapiszewski. 2014. "Data Access and Research Transparency in the Qualitative Tradition." *PS: Political Science and Politics* vol. 47, no. 1: 43–47.

Fairfield, Tasha. 2013. "Going Where the Money Is: Strategies for Taxing Economic Elites in Unequal Democracies." *World Development* vol. 47 (July): 42–57.

———. 2015. "Reflections on Analytic Transparency in Process Tracing Research." *Qualitative and Multi-Method Research: Newsletter of the American Political Science Association's QMMR Section* vol. 13, no. 1: 46–50.

Fairfield, Tasha, and Charman, Andrew. 2015. "Bayesian Probability: The Logic of (Political) Science." Paper prepared for the Annual Meeting of the American Political Science Association, Sept. 3-6, San Francisco.

Fearon, James D., and David D. Laitin. 2003. "Ethnicity, Insurgency, and Civil War." *American Political Science Review* vol. 97, no. 1: 75–90.

Fenno, Richard F. 1986. "Observation, Context, and Sequence in the Study of Politics: Presidential Address Presented at the 81st Annual Meeting of the APSA, New Orleans, 29 August 1985." *American Political Science Review* vol. 80, no. 1: 3–15.

Finnemore, Martha. 2013. "Constructing Statistics for Global Governance." Unpublished manuscript, George Washington University.

Fiorina, Morris P. 1995. "Rational Choice, Empirical Contributions, and the Scientific Enterprise." *Critical Review* vol. 9, no. 1-2: 85–94.

Gallie, W. B. 1955-1956. "Essentially Contested Concepts." *Proceedings of the Aristotelian Society* vol. 56 (1955-1956): 167–198.

Golub, Stephen S., Céline Kauffmann, and Philip Yeres. 2011. "Defining and Measuring Green FDI: An Exploratory Review of Existing Work and Evidence." OECD Working Papers on International Investment no. 2011/02. Online at http://dx.doi.org/10.1787/5kg58j1 cvcvk-en.

Herrera, Yoshiko M., and Devesh Kapur. 2007. "Improving Data Quality: Actors, Incentives, and Capabilities." *Political Analysis* vol. 15, no. 4: 365–386.

Henson, Spencer, and Rupert Loader. 2001. "Barriers to Agricultural Exports from Developing Countries: The Role of Sanitary and Phytosanitary Requirements." *World Development* vol. 29, no. 1: 85–102.

Humphreys, Macartan, and Alan M. Jacobs. Forthcoming. "Mixing Methods: A Bayesian Approach." *American Political Science Review*.

Humphreys, Macartan, Raul Sanchez de la Sierra, Peter van der Windt. 2013. "Fishing, Commitment, and Communication: A Proposal for Comprehensive Nonbinding Research Registration." *Political Analysis* vol. 21, no. 1: 1–20.

International Monetary Fund (IMF). Various years. *Foreign Direct Investment Statistics: How Countries Measure FDI*. Washington D.C.: IMF, 2001, 2003, etc.

Isaac, Jeffrey C. 2015. "For a More Public Political Science." *Perspectives on Politics* vol. 13 no. 2: 269–283.

Jackson, Patrick Thaddeus. 2008. "Foregrounding Ontology: Dualism, Monism, and IR Theory." *Review of International Studies* vol. 34, no. 1: 129–153.

Jervis, Robert. 1976. *Perception and Misperception in International Politics*. Princeton: Princeton University Press.

———. 1997. *System Effects: Complexity in Political and Social Life*. Princeton: Princeton University Press.

Katz, Michael L., and Carl Shapiro. 1985. "Network Externalities, Competition, and Compatibility." *American Economic Review* vol. 75, no. 3: 424–440.

King, Gary. 1995. "Replication, Replication." *PS: Political Science and Politics* vol. 28, no. 3: 444–452.

King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry*. Princeton: Princeton University Press.

Lauterbach, Claire. 2007. "The Costs of Cooperation: Civilian Casualty Counts in Iraq." *International Studies Perspectives* vol. 8, no. 4: 429–445.

Lawrence, Adria. 2010. "Triggering Nationalist Violence Competition and Conflict in Uprisings against Colonial Rule." *International Security* vol. 35, no. 2: 88–122.

Lukes, Steven. 1974. *Power: A Radical View*, Studies in Sociology. London: MacMillan.

Lupia, Arthur, and Colin Elman. 2014. "Introduction: Openness in Political Science: Data Access and Transparency." *PS: Political Science and Politics* vol. 47, no. 1: 19–23.

Maskus, Keith E., Tsunehiro Otsuki, and John S. Wilson. 2005. "The Cost of Compliance with Product Standards for Firms in Developing Countries: An Econometric Study." *World Bank Policy Research Paper* no.3590 (May).

Moravcsik, Andrew. 2010. "Active Citation: A Precondition for Replicable Qualitative Research." *PS: Political Science and Politics* vol. 43, no. 1: 29–35.

Nye, Joseph. 2009. "Scholars on the Sidelines." *Washington Post* 13 April.

Pachirat, Timothy. 2015. "The Tyranny of Light." *Qualitative and Multi-Method Research: Newsletter of the American Political Science Association's QMMR Section* vol. 13, no. 1: 27–31.

Parkinson, Sarah Elizabeth, and Elisabeth Jean Wood. 2015. "Transparency in Intensive Research on Violence: Ethical Dilemmas and Unforeseen Consequences." *Qualitative and Multi-Method Research: Newsletter of the American Political Science Association's QMMR Section* vol. 13, no. 1: 22–27.

Phillips, Trisha and Frachesca Nestor. 2015. "Ethical Standards in Political Science: Journals, Articles, and Education." Unpublished manuscript, West Virginia University.

Pierson, Paul, and Theda Skocpol. 2002. "Historical Institutionalism in Contemporary Political Science." In *Political Science: The State of the Discipline*, edited by Ira Katznelson and Helen V. Milner. New York: W. W. Norton for the American Political Science Association: 693–721.

Putnam, Robert D. 2003. "The Public Role of Political Science." *Perspectives on Politics* vol.1 no.2: 249–255.

Rege, Vinod, Shyam K. Gujadhur, and Roswitha Franz. 2003. *Influ-*

encing and Meeting International Standards: Challenges for Developing Countries. Geneva/London: UNCTAD/WTO International Trade Center and Commonwealth Secretariat.

Resnik, David B., Elizabeth Wager, and Grace E. Kissling. 2015. "Retraction Policies of Top Scientific Journals Ranked by Impact Factor." *Journal of the Medical Library Association* vol.103 (forthcoming).

Rogowski, Ronald. 1995. "The Role of Theory and Anomaly in Social-Scientific Inference." *American Political Science Review* vol. 89, no. 2: 467–470.

Romney, David, Brandon M. Stewart, and Dustin Tingley. 2015. "Plain Text? Transparency in Computer-Assisted Text Analysis." *Qualitative and Multi-Method Research: Newsletter of the American Political Science Association's QMMR Section* vol. 13, no. 1: 32–37.

Saey, Tina Hesman. 2015. "Repeat Performance: Too Many Studies, When Replicated, Fail to Pass Muster." *Science News* vol. 187, no. 2 (24 January): 21–26.

Sambanis, Nicholas. 2004. "What Is Civil War? Conceptual and Empirical Complexities of an Operational Definition." *Journal of Conflict Resolution* vol. 48, no. 6: 814–858.

Saunders, Elizabeth N. 2014. "Transparency without Tears: A Pragmatic Approach to Transparent Security Studies Research." *Security Studies* vol. 23, no. 4: 689–698.

Shih, Victor. 2015. "Research in Authoritarian Regimes: Transparency Tradeoffs and Solutions." *Qualitative and Multi-Method Research: Newsletter of the American Political Science Association's QMMR Section* vol. 13, no. 1: 20–22.

Smith, Richard. 2008. "Most Cases of Research Misconduct Go Undetected, Conference Told." *British Medical Journal* vol. 336, no. 7650 (26 April 2008): 913.

Sniderman, Paul M. 1995. "Evaluation Standards for a Slow-Moving Science." *PS: Political Science and Politics* vol. 28, no. 3: 464–467.

Snyder, Jack. 1984. *The Ideology of the Offensive: Military Decision Making and the Disasters of 1914*. Ithaca, NY: Cornell University Press.

———. 2014. "Active Citation: In Search of Smoking Guns or Meaningful Context?" *Security Studies* vol. 23, no. 4: 708–714.

———. 2015. "Russia: The Politics and Psychology of Overcommitment." (First published in *The Ideology of the Offensive: Military Decision Making and the Disasters of 1914*. Ithaca, NY: Cornell University Press, 1984: 165–198).) Active Citation Compilation, QDR:10047. Syracuse, NY: Qualitative Data Repository [distributor]. Online at https://qdr.syr.edu/discover/projectdescriptionssnyder (last accessed 8/5/2015).

Tetlock, Philip E. 2005. *Expert Political Judgment: How Good Is It? How Can We Know?* Princeton: Princeton University Press.

Trachtenberg, Marc. 2015. "Transparency in Practice: Using Written Sources." *Qualitative and Multi-Method Research: Newsletter of the American Political Science Association's QMMR Section* vol. 13, no. 1: 13–17.

Wagemann, Claudius, and Carsten Q. Schneider. 2015. "Transparency Standards in Qualitative Comparative Analysis." *Qualitative and Multi-Method Research: Newsletter of the American Political Science Association's QMMR Section* vol. 13, no. 1: 38–42.

Wager, Elizabeth. 2015. "Why Are Retractions So Difficult." *Science Editing* vol. 2, no. 1: 32–34.

Qualitative and Multi-Method Research Journal Scan:
January 2014 – April 2015

In this Journal Scan, we provide citations and (where available) abstracts of articles that have appeared in political science and related journals from January 2014 through April 2015 and which address some facet of qualitative methodology or multi-method research. The Journal Scan's focus is on articles that develop an explicitly methodological argument or insight; it does not seek to encompass applications of qualitative or multiple methods. A list of all journals consulted is provided at the end of this section.†

Analysis of Text

D'Orazio, Vito, Steven T. Landis, Glenn Palmer, and Philip Schrodt. 2014. "Separating the Wheat from the Chaff: Applications of Automated Document Classification Using Support Vector Machines." *Political Analysis* vol. 22, no. 2: 224–242. DOI: 10.1093/pan/mpt030

Due in large part to the proliferation of digitized text, much of it available for little or no cost from the Internet, political science research has experienced a substantial increase in the number of data sets and large-n research initiatives. As the ability to collect detailed information on events of interest expands, so does the need to efficiently sort through the volumes of available information. Automated document classification presents a particularly attractive methodology for accomplishing this task. It is efficient, widely applicable to a variety of data collection efforts, and considerably flexible in tailoring its application for specific research needs. This article offers a holistic review of the application of automated document classification for data collection in political science research by discussing the process in its entirety. We argue that the application of a two-stage support vector machine (SVM) classification process offers advantages over other well-known alternatives, due to the nature of SVMs being a discriminative classifier and having the ability to effectively address two primary attributes of textual data: high dimensionality and extreme sparseness. Evidence for this claim is presented through a discussion of the efficiency gains derived from using automated document classification on the Militarized Interstate Dispute 4 (MID4) data collection project.

Lauderdale, Benjamin E., and Tom S. Clark. 2014. "Scaling Politically Meaningful Dimensions Using Texts and Votes." *American Journal of Political Science* vol. 58, no. 3: 754–771. DOI: 10.1111/ajps.12085

Item response theory models for roll-call voting data provide political scientists with parsimonious descriptions of political actors' relative preferences. However, models using only voting data tend to obscure variation in preferences across different issues due to identification and labeling problems that

arise in multidimensional scaling models. We propose a new approach to using sources of metadata about votes to estimate the degree to which those votes are about common issues. We demonstrate our approach with votes and opinion texts from the U.S. Supreme Court, using latent Dirichlet allocation to discover the extent to which different issues were at stake in different cases and estimating justice preferences within each of those issues. This approach can be applied using a variety of unsupervised and supervised topic models for text, community detection models for networks, or any other tool capable of generating discrete or mixture categorization of subject matter from relevant vote-specific metadata.

Laver, Michael. 2014. "Measuring Policy Positions in Political Space." *Annual Review of Political Science* vol. 17, no. 1: 207–223. DOI: doi:10.1146/annurev-polisci-061413-041905

Spatial models are ubiquitous within political science. Whenever we confront spatial models with data, we need valid and reliable ways to measure policy positions in political space. I first review a range of general issues that must be resolved before thinking about how to measure policy positions, including cognitive metrics, a priori and a posteriori scale interpretation, dimensionality, common spaces, and comparability across settings. I then briefly review different types of data we can use to do this and measurement techniques associated with each type, focusing on headline issues with each type of data and pointing to comprehensive surveys of relevant literatures—including expert, elite, and mass surveys; text analysis; and legislative voting behavior.

Lucas, Christopher, Richard A. Nielsen, Margaret E. Roberts, Brandon M. Stewart, Alex Storer, and Dustin Tingley. 2015. "Computer-Assisted Text Analysis for Comparative Politics." *Political Analysis* vol. 23, no. 2: 254–277. DOI: 10.1093/pan/mpu019

Recent advances in research tools for the systematic analysis of textual data are enabling exciting new research throughout the social sciences. For comparative politics, scholars who are often interested in non-English and possibly multilingual textual datasets, these advances may be difficult to access. This article discusses practical issues that arise in the processing, management, translation, and analysis of textual data with a particular focus on how procedures differ across languages.

These procedures are combined in two applied examples of automated text analysis using the recently introduced Structural Topic Model. We also show how the model can be used to analyze data that have been translated into a single language via machine translation tools. All the methods we describe here are implemented in open-source software packages available from the authors.

Roberts, Margaret E., Brandon M. Stewart, Dustin Tingley, Christopher Lucas, Jetson Leder-Luis, Shana Kushner Gadarian, Bethany Albertson, and David G. Rand. 2014. "Structural Topic Models for Open-Ended Survey Responses." *American Journal of Political Science* vol. 58, no. 4: 1064–1082. DOI: 10.1111/ajps.12103

Collection and especially analysis of open-ended survey responses are relatively rare in the discipline and when conducted are almost exclusively done through human coding. We present an alternative, semiautomated approach, the structural topic model (STM) (Roberts, Stewart, and Airoldi 2013; Roberts et al. 2013), that draws on recent developments in machine learning based analysis of textual data. A crucial contribution of the method is that it incorporates information about the document, such as the author's gender, political affiliation, and treatment assignment (if an experimental study). This article focuses on how the STM is helpful for survey researchers and experimentalists. The STM makes analyzing open-ended responses easier, more revealing, and capable of being used to estimate treatment effects. We illustrate these innovations with analysis of text from surveys and experiments.

Case Selection and Sampling

Roll, Kate. 2014. "Encountering Resistance: Qualitative Insights from the Quantitative Sampling of Ex-Combatants in Timor-Leste." *PS: Political Science & Politics* vol. 47, no. 2: 485–489. DOI: doi:10.1017/S1049096514000420

This article highlights the contribution of randomized, quantitative sampling techniques to answering qualitative questions posed by the study. In short it asks: what qualitative insights do we derive from quantitative sampling processes? Rather than simply being a means to an end, I argue the sampling process itself generated data. More specifically, seeking out more than 220 geographically dispersed individuals, selected though a randomized cluster sample, resulted in the identification of relationship patterns, highlighted extant resistance-era hierarchies and patronage networks, as well as necessitated deeper, critical engagement with the sampling framework. While this discussion is focused on the study of former resistance members in Timor-Leste, these methodological insights are broadly relevant to researchers using mixed methods to study former combatants or other networked social movements.

Case Studies and Comparative Method

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2015. "Comparative Politics and the Synthetic Control Method." *American Journal of Political Science* vol. 59, no. 2: 495–510. DOI: 10.1111/ajps.12116

In recent years, a widespread consensus has emerged about the necessity of establishing bridges between quantitative and qualitative approaches to empirical research in political science. In this article, we discuss the use of the synthetic control method as a way to bridge the quantitative/qualitative divide in comparative politics. The synthetic control method provides a systematic way to choose comparison units in comparative case studies. This systematization opens the door to precise quantitative inference in small-sample comparative studies, without precluding the application of qualitative approaches. Borrowing the expression from Sidney Tarrow, the synthetic control method allows researchers to put "qualitative flesh on quantitative bones." We illustrate the main ideas behind the synthetic control method by estimating the economic impact of the 1990 German reunification on West Germany.

Braumoeller, Bear F. 2014. "Information and Uncertainty: Inference in Qualitative Case Studies." *International Studies Quarterly* vol. 58, no. 4: 873-875. DOI: 10.1111/isqu.12169

Drozdova and Gaubatz (2014) represent a welcome addition to the growing literature on quantitative methods designed to complement qualitative case studies. Partly due to its crossover nature, however, the article balances delicately—and ultimately untenably—between within-sample and out-of-sample inference. Moreover, isomorphisms with existing techniques, while validating the methodology, simultaneously raise questions regarding its comparative advantage.

Drozdova, Katya, and Kurt Taylor Gaubatz. 2014. "Reducing Uncertainty: Information Analysis for Comparative Case Studies." *International Studies Quarterly* vol. 58, no. 3: 633–645. DOI: 10.1111/isqu.12101

The increasing integration of qualitative and quantitative analysis has largely focused on the benefits of in-depth case studies for enhancing our understanding of statistical results. This article goes in the other direction to show how some very straightforward quantitative methods drawn from information theory can strengthen comparative case studies. Using several prominent "structured, focused comparison" studies, we apply the information-theoretic approach to further advance these studies' findings by providing systematic, comparable, and replicable measures of uncertainty and influence for the factors they identified. The proposed analytic tools are simple enough to be used by a wide range of scholars to enhance comparative case study findings and ensure the maximum leverage for discerning between alternative explanations as well as cumulating knowledge from multiple studies. Our approach especially serves qualitative policy-relevant case comparisons in international studies, which have typically avoided more complex or less applicable quantitative tools.

Gisselquist, Rachel M. 2014. "Paired Comparison and Theory Development: Considerations for Case Selection." *PS: Political Science & Politics* vol. 47, no. 2: 477–484. DOI: doi:10.1017/S1049096514000419

Despite the widespread use of paired comparisons, we lack clear guidance about how to use this research strategy in practice, particularly in case selection. The literature tends to assume that cases are systematically selected from a known population, a major assumption for many topics of interest to political scientists. This article speaks to this gap. It describes three distinct logics of paired comparison relevant to theory development, presents a simple way of considering and comparing them, and explores how this approach can inform more intentional research design, with particular attention to low information settings where substantial research is needed to ascertain the values of independent or dependent variables. The discussion underscores inter alia the need to be aware and explicit about the implications of case selection for the ability to test and build theory, and the need to reconsider the well-cited "rule" of not selecting on the dependent variable.

Pieczara, Kamila, and Yong-Soo Eun. 2014. "Smoke, but No Fire? In Social Science, Focus on the Most Distinct Part." *PS: Political Science & Politics* vol. 47, no. 1: 145–148. DOI: doi:10.1017/S104909651300156X

Causality in social science is hard to establish even through the finest comparative research. To ease the task of extracting causes from comparisons, we present the benefits of tracing particularities in any phenomenon under investigation. We introduce three real-world examples from 2011: British riots, worldwide anticapitalist protests, and the highway crash near Taunton in southwestern England. Whereas all of these three examples have broad causes, we embark on the quest after specific factors. The Taunton accident can send a powerful message to social scientists, which is about the danger of making general statements in their explanations. Instead of saying much but explaining little, the merit of singling out the specific is substantial. As social scientists, when we are faced with "smoke" but no "fire," let us then focus on the part that is distinct.

Rohlfing, Ingo. 2014. "Comparative Hypothesis Testing Via Process Tracing." *Sociological Methods & Research* vol. 43, no. 4: 606–642. DOI: 10.1177/0049124113503142

Causal inference via process tracing has received increasing attention during recent years. A 2 × 2 typology of hypothesis tests takes a central place in this debate. A discussion of the typology demonstrates that its role for causal inference can be improved further in three respects. First, the aim of this article is to formulate case selection principles for each of the four tests. Second, in focusing on the dimension of uniqueness of the 2 × 2 typology, I show that it is important to distinguish between theoretical and empirical uniqueness when choosing cases and generating inferences via process tracing. Third, I demonstrate that the standard reading of the so-called doubly decisive test is misleading. It conflates unique implications of a hypothesis with contradictory implications between one hypothesis and another. In order to remedy the current ambiguity of the dimension of uniqueness, I propose an expanded typology of hypothesis tests that is constituted by three dimensions.

Thiem, Alrik. 2014. "Unifying Configurational Comparative Methods: Generalized-Set Qualitative Comparative Analysis." *Sociological Methods & Research* vol. 43, no. 2: 313–337. DOI: 10.1177/0049124113500481

Crisp-set Qualitative Comparative Analysis, fuzzy-set Qualitative Comparative Analysis (fsQCA), and multi-value Qualitative Comparative Analysis (mvQCA) have emerged as distinct variants of QCA, with the latter still being regarded as a technique of doubtful set-theoretic status. Textbooks on configurational comparative methods have emphasized differences rather than commonalities between these variants. This article has two consecutive objectives, both of which focus on commonalities. First, but secondary in importance, it demonstrates that all set types associated with each variant can be combined within the same analysis by introducing a standardized notational system. By implication, any doubts about the set-theoretic status of mvQCA vis-à-vis its two sister variants are removed. Second, but primary in importance and dependent on the first objective, this article introduces the concept of the multivalent fuzzy set variable. This variable type forms the basis of generalized-set Qualitative Comparative Analysis (gsQCA), an approach that integrates the features peculiar to mvQCA and fsQCA into a single framework while retaining routine truth table construction and minimization procedures. Under the concept of the multivalent fuzzy set variable, all existing QCA variants become special cases of gsQCA.

Causality and Causal Inference

Scholarly Exchange: "Effects of Causes and Causes of Effects"

Dawid, A. Philip, David L. Faigman, and Stephen E. Fienberg. 2014. "Fitting Science Into Legal Contexts: Assessing Effects of Causes or Causes of Effects?" *Sociological Methods & Research* vol. 43, no. 3: 359–390. DOI: 10.1177/0049124113515188

Law and science share many perspectives, but they also differ in important ways. While much of science is concerned with the effects of causes (EoC), relying upon evidence accumulated from randomized controlled experiments and observational studies, the problem of inferring the causes of effects (CoE) requires its own framing and possibly different data. Philosophers have written about the need to distinguish between the "EoC" and "the CoE" for hundreds of years, but their advice remains murky even today. The statistical literature is only of limited help here as well, focusing largely on the traditional problem of the "EoC." Through a series of examples, we

review the two concepts, how they are related, and how they differ. We provide an alternative framing of the "CoE" that differs substantially from that found in the bulk of the scientific literature, and in legal cases and commentary on them. Although in these few pages we cannot fully resolve this issue, we hope to begin to sketch a blueprint for a solution. In so doing, we consider how causation is framed by courts and thought about by philosophers and scientists. We also endeavor to examine how law and science might better align their approaches to causation so that, in particular, courts can take better advantage of scientific expertise.

Jewell, Nicholas P. 2014. "Assessing Causes for Individuals: Comments on Dawid, Faigman, and Fienberg." *Sociological Methods & Research* vol. 43, no. 3: 391–395. DOI: 10.1177/0049124113518190

In commenting on Dawid, Faigman, and Fienberg, the author contrasts the proposed parameter, the probability of causation, to other parameters in the causal inference literature, specifically the probability of necessity discussed by Pearl, and Robins and Greenland, and Pearl's probability of sufficiency. This article closes with a few comments about the difficulties of estimation of parameters related to individual causation.

Cheng, Edward K. 2014. "Comment on Dawid, Faigman, and Fienberg (2014)." *Sociological Methods & Research* vol. 43, no. 3: 396–400. DOI: 10.1177/0049124113518192

Beecher-Monas, Erica. 2014. "Comment on Philip Dawid, David Faigman, and Stephen Fienberg, Fitting Science into Legal Contexts: Assessing Effects of Causes or Causes of Effects." *Sociological Methods & Research* vol. 43, no. 3: 401–405. DOI: 10.1177/0049124113518191

Smith, Herbert L. 2014. "Effects of Causes and Causes of Effects: Some Remarks From the Sociological Side." *Sociological Methods & Research* vol. 43, no. 3: 406–415. DOI: 10.1177/0049124114521149

Sociology is pluralist in subject matter, theory, and method, and thus a good place to entertain ideas about causation associated with their use under the law. I focus on two themes: (1) the legal lens on causation that "considers populations in order to make statements about individuals" and (2) the importance of distinguishing between effects of causes and causes of effects.

Dawid, A. Philip, David L. Faigman, and Stephen E. Fienberg. 2014. "Authors' Response to Comments on Fitting Science Into Legal Contexts: Assessing Effects of Causes or Causes of Effects?" *Sociological Methods & Research* vol. 43, no. 3: 416–421. DOI: 10.1177/0049124113515189

Dawid, A. Philip, David L. Faigman, and Stephen E. Fienberg. 2015. "On the Causes of Effects: Response to Pearl." *Sociological Methods & Research* vol. 44, no. 1: 165–174. DOI: 10.1177/0049124114562613

We welcome Professor Pearl's comment on our original article, Dawid et al. Our focus there on the distinction between the "Effects of Causes" (EoC) and the "Causes of Effects" (CoE) concerned two fundamental problems, one a theoretical challenge in statistics and the other a practical challenge for trial courts. In this response, we seek to accomplish several things. First, using Pearl's own notation, we attempt to clarify the similarities and differences between his technical approach and that in Dawid et al. Second, we consider the more practical challenges for CoE in the trial court setting, and explain why we believe Pearl's analyses, as described via his example, fail to address these. Finally, we offer some concluding remarks.

Pearl, Judea. 2015. "Causes of Effects and Effects of Causes." *Sociological Methods & Research* vol. 44, no. 1: 149–164. DOI: 10.1177/0049124114562614

This article summarizes a conceptual framework and simple mathematical methods of estimating the probability that one event was a necessary cause of another, as interpreted by lawmakers. We show that the fusion of observational and experimental data can yield informative bounds that, under certain circumstances, meet legal criteria of causation. We further investigate the circumstances under which such bounds can emerge, and the philosophical dilemma associated with determining individual cases from statistical data.

Communication of Research and Results

Gelman, Andrew, and Thomas Basbøll. 2014. "When Do Stories Work? Evidence and Illustration in the Social Sciences." *Sociological Methods & Research* vol. 43, no. 4: 547–570. DOI: 10.1177/0049124114526377

Storytelling has long been recognized as central to human cognition and communication. Here we explore a more active role of stories in social science research, not merely to illustrate concepts but also to develop new ideas and evaluate hypotheses, for example, in deciding that a research method is effective. We see stories as central to engagement with the development and evaluation of theories, and we argue that for a story to be useful in this way, it should be anomalous (representing aspects of life that are not well explained by existing models) and immutable (with details that are well-enough established that they have the potential to indicate problems with a new model). We develop these ideas through considering two well-known examples from the work of Karl Weick and Robert Axelrod, and we discuss why transparent sourcing (in the case of Axelrod) makes a story a more effective research tool, whereas improper sourcing (in the case of Weick) interferes with the key useful roles of stories in the scientific process.

Mahoney, James, and Rachel Sweet Vanderpoel. 2015. "Set Diagrams and Qualitative Research." *Comparative Political Studies* vol. 48, no. 1: 65–100. DOI: 10.1177/0010414013519410

Political scientists have developed important new ideas for using spatial diagrams to enhance quantitative research. Yet the potential uses of diagrams for qualitative research have not been explored systematically. We begin to correct this omission by showing how set diagrams can facilitate the application of qualitative methods and improve the presentation of qualitative findings. Set diagrams can be used in conjunction with a wide range of qualitative methodologies, including process tracing, concept formation, counterfactual analysis, sequence elaboration, and qualitative comparative analysis. We illustrate the utility of set diagrams by drawing on substantive examples of qualitative research in the fields of international relations and comparative politics.

Critical Theory as Empirical Methodology

Patberg, Markus. 2014. "Supranational Constitutional Politics and the Method of Rational Reconstruction." *Philosophy & Social Criticism* vol. 40, no. 6: 501–521. DOI: 10.1177/0191453714530987

In The Crisis of the European Union Jürgen Habermas claims that the constituent power in the EU is shared between the community of EU citizens and the political communities of the member states. By his own account, Habermas arrives at this concept of a dual constituent subject through a rational reconstruction of the genesis of the European constitution. This explanation, however, is not particularly illuminating since it is controversial what the term 'rational reconstruction' stands for. This article critically discusses the current state of research on rational reconstruction, develops a new reading of Habermas' method and invokes this account for an explanation and evaluation of the notion of a European pouvoir constituant mixte.

Ethnography

Scholarly Exchange: "Ethnography and the Attitudinal Fallacy."

Jerolmack, Colin, and Shamus Khan. 2014. "Talk Is Cheap: Ethnography and the Attitudinal Fallacy." *Sociological Methods & Research* vol. 43, no. 2: 178–209. DOI: 10.1177/0049124114523396

This article examines the methodological implications of the fact that what people say is often a poor predictor of what they do. We argue that many interview and survey researchers routinely conflate self-reports with behavior and assume a consistency between attitudes and action. We call this erroneous inference of situated behavior from verbal accounts the attitudinal fallacy. Though interviewing and ethnography are often lumped together as "qualitative methods," by juxtaposing studies of "culture in action" based on verbal accounts with ethnographic investigations, we show that the latter routinely attempts to explain the "attitude–behavior problem" while the former regularly ignores it. Because meaning and action are collectively negotiated and context-dependent, we contend that self-reports of attitudes and behaviors are of limited value in explaining what people actually do because they are overly individualistic and abstracted from lived experience.

Maynard, Douglas W. 2014. "News From Somewhere, News From Nowhere : On the Study of Interaction in Ethnographic Inquiry." *Sociological Methods & Research* vol. 43, no. 2: 210–218. DOI: 10.1177/0049124114527249

This is a comment suggesting that Jerolmack and Khan's article in this issue embodies news from "somewhere," in arguing that ethnography can emphasize interaction in concrete situations and what people do rather than what they say about what they do. However, their article also provides news from "nowhere," in that ethnography often claims to prioritize in situ organization while dipping into an unconstrained reservoir of distant structures that analytically can subsume and potentially eviscerate the local order. I elaborate on each of these somewhere/nowhere ideas. I also briefly point to the considerable ethnomethodological and conversation analytic research of the last several decades that addresses the structural issue. Such research, along with other traditions in ethnography, suggest that investigators can relate social or political contexts to concrete situations provided that there is, in the first place, preservation of the parameters of everyday life and the exactitude of the local order.

Cerulo, Karen A. 2014. "Reassessing the Problem: Response to Jerolmack and Khan." *Sociological Methods & Research* vol. 43, no. 2: 219–226. DOI: 10.1177/0049124114526378

This article offers reflections on Jerolmack and Khan's article "Talk is Cheap: Ethnography and the Attitudinal Fallacy." Specifically, I offer three suggestions aimed at moderating the authors' critique. Since the sociology of culture and cognition is my area of expertise, I, like Jerolmack and Khan, use this literature to mine supporting examples.

Vaisey, Stephen. 2014. "The "Attitudinal Fallacy" Is a Fallacy: Why We Need Many Methods to Study Culture." *Sociological Methods & Research* vol. 43, no. 2: 227–231. DOI: 10.1177/0049124114523395

DiMaggio, Paul. 2014. "Comment on Jerolmack and Khan, "Talk Is Cheap": Ethnography and the Attitudinal Fallacy." *Sociological Methods & Research* vol. 43, no. 2: 232–235. DOI: 10.1177/0049124114526371

Jerolmack, Colin, and Shamus Khan. 2014. "Toward an Understanding of the Relationship Between Accounts and Action." *Sociological Methods & Research* vol. 43, no. 2: 236–247. DOI: 10.1177/0049124114523397

Katz, Jack. 2015. "Situational Evidence: Strategies for Causal Reasoning From Observational Field Notes." *Sociological Methods & Research* vol. 44, no. 1: 108–44. DOI: 10.1177/0049124114554870

There is unexamined potential for developing and testing rival causal explanations in the type of data that participant observation is best suited to create: descriptions of in situ social interaction crafted from the participants' perspectives. By intensively examining a single ethnography, we can see how multiple predictions can be derived from and tested with field notes, how numerous strategies are available for demonstrating the patterns of nonoccurrence which causal propositions imply, how qualitative data can be analyzed to negate researcher behavior as an alternative causal explanation, and how folk counterfactuals can add to the evidentiary strength of an ethnographic study. Explicating the potential of field notes for causal explanation may be of interest to methodologists who seek a common logic for guiding and evaluating quantitative and qualitative research, to ethnographic fieldworkers who aim at connecting micro- to macro-social processes, to researchers who use an analogous logic of explanation when working with other forms of qualitative data, and to comparative–analytic sociologists who wish to form concepts and develop theory in conformity with an understanding that social life consists of social interaction processes that may be captured most directly by ethnographic fieldwork.

Field Research

Symposium: "Fieldwork in Political Science: Encountering Challenges and Crafting Solutions."

Hsueh, Roselyn, Francesca Refsum Jensenius, and Akasemi Newsome. 2014. "Introduction." *PS: Political Science & Politics* vol. 47, no. 2: 391–393. DOI: doi:10.1017/S1049096514000262

Whether the aim is to build theory or test hypotheses, junior and senior political scientists alike face problems collecting data in the field. Most field researchers have expectations of the challenges they will face, and also some training and preparation for addressing these challenges. Yet, in hindsight many wish they had been better prepared—both psychologically and logistically—for the difficulties they encountered. The central theme of this symposium is precisely these data collection problems political scientists face in the field and how to deal with them. The separate perspectives presented here contextualize particular challenges of data collection in different world regions within the trajectory of single research projects. The articles trace the challenges that analysts faced in field sites as varied as China, Germany, India, Kazakhstan, and Mexico. Describing the realities of fieldwork and resourceful strategies for dealing with them, this symposium sheds new light on several practical aspects of fieldwork in political science. The symposium also brings together scholars who used multiple research methods, thereby illuminating the difficulties encountered in political science fieldwork from diverse angles. For this reason, these vignettes are relevant to researchers focusing on both qualitative and quantitative research methods.

Scoggins, Suzanne E. 2014. "Navigating Fieldwork as an Outsider: Observations from Interviewing Police Officers in China." *PS: Political Science & Politics* vol. 47, no. 2: 394–397. DOI: doi:10.1017/S1049096514000274

Sirnate, Vasundhara. 2014. "Positionality, Personal Insecurity, and Female Empathy in Security Studies Research." *PS: Political Science & Politics* vol. 47, no. 2: 398–401. DOI: doi:10.1017/S1049096514000286

Jensenius, Francesca Refsum. 2014. "The Fieldwork of Quantitative Data Collection." *PS: Political Science & Politics* vol. 47, no. 2: 402–404. DOI: doi:10.1017/S1049096514000298

Chambers-Ju, Christopher. 2014. "Data Collection, Opportunity Costs, and Problem Solving: Lessons from Field Research on Teachers' Unions in Latin America." *PS: Political Science & Politics* vol. 47, no. 2: 405–409. DOI: doi:10.1017/S1049096514000304

Newsome, Akasemi. 2014. "Knowing When to Scale Back: Addressing Questions of Research Scope in the Field." *PS: Political Science & Politics* vol. 47, no. 2: 410–413. DOI: doi:10.1017/S1049096514000316

LaPorte, Jody. 2014. "Confronting a Crisis of Research Design." *PS: Political Science & Politics* vol. 47, no. 2: 414–417. DOI: doi:10.1017/S1049096514000328

Interpretive and Hermeneutic Approaches

Epstein, Charlotte. 2015. "Minding the Brain: IR as a Science?" *Millennium - Journal of International Studies* vol. 43, no. 2: 743–748. DOI: 10.1177/0305829814557558

Invited by the editors to respond to Professor Neumann's inaugural lecture, in this article I take issue with his core, unquestioned assumption, namely, whether IR should be considered as a science. I use it as a starting point to re-open the question of how the stuff that humans are made of should be studied in IR today. Beyond Neumann's piece, I critically engage with two emerging trends in the discipline, the so-called new materialisms and the interest in the neurosciences, and articulate my concern that these trends have not addressed the deterministic fallacy that threatens to undermine their relevance for the study of a world made by humans. To the latent anxiety as to whether the discipline has finally achieved recognition of its epistemological status as a science, I respond by recalling that other grand tradition in IR, interpretive methods. The study of meaning from within, without reducing it to countable 'things' or to neuronal traces, is, I suggest, better attuned to capturing the contingency, indeterminacy and freedom which constitute key characteristics of the constructed, social world that we study in IR.

Ginev, Dimitri. 2014. "Radical Reflexivity and Hermeneutic Pre-normativity." *Philosophy & Social Criticism* vol. 40, no. 7: 683–703. DOI: 10.1177/0191453714536432

This article develops the thesis that normative social orders are always fore-structured by horizons of possibilities. The thesis is spelled out against the background of a criticism of ethnomethodology for its hermeneutic deficiency in coping with radical reflexivity. The article contributes to the debates concerning the status of normativity problematic in the cultural disciplines. The concept of hermeneutic pre-normativity is introduced to connote the interpretative fore-structuring of normative inter-subjectivity. Radical reflexivity is reformulated in terms of hermeneutic phenomenology.

Measurement and Concept Formation

Seawright, Jason, and David Collier. 2014. "Rival Strategies of Validation: Tools for Evaluating Measures of Democracy." *Comparative Political Studies* vol. 47, no. 1: 111–138. DOI: 10.1177/0010414013489098

The challenge of finding appropriate tools for measurement validation is an abiding concern in political science. This article considers four traditions of validation, using examples from cross-national research on democracy: the levels-of-measurement approach, structural-equation modeling with latent variables, the pragmatic tradition, and the case-based method. Methodologists have sharply disputed the merits of alternative traditions. We encourage scholars—and certainly analysts of democracy—to pay more attention to these disputes and to consider strengths and weaknesses in the validation tools they adopt. An online appendix summarizes the evaluation of six democracy data sets from the perspective of alternative approaches to validation. The overall goal is to open a new discussion of alternative validation strategies.

Wilson, Matthew C. 2014. "A Discreet Critique of Discrete Regime Type Data." *Comparative Political Studies* vol. 47, no. 5: 689–714. DOI: 10.1177/0010414013488546

To understand the limitations of discrete regime type data for studying authoritarianism, I scrutinize three regime type data sets provided by Cheibub, Gandhi, and Vreeland, Hadenius and Teorell, and Geddes. The political narratives of Nicaragua, Colombia, and Brazil show that the different data sets on regime type lend themselves to concept stretching and misuse, which threatens measurement validity. In an extension of Fjelde's analysis of civil conflict onset, I demonstrate that interchangeably using the data sets leads to divergent predictions, it is sensitive to outliers, and the data ignore certain institutions. The critique expounds on special issues with discrete data on regime type so that scholars make more informed choices and are better able to compare results. The mixed-methods assessment of discrete data on regime type demonstrates the importance of proper concept formation in theory testing. Maximizing the impact of such data requires the scholar to make more theoretically informed choices.

Yom, Sean. 2015. "From Methodology to Practice: Inductive Iteration in Comparative Research." *Comparative Political Studies* vol. 48, no. 5: 616–644. DOI: 10.1177/0010414014554685

Most methods in comparative politics prescribe a deductive template of research practices that begins with proposing hypotheses, proceeds into analyzing data, and finally concludes with confirmatory tests. In reality, many scholars move back and forth between theory and data in creating causal explanations, beginning not with hypotheses but hunches and constantly revising their propositions in response to unexpected discoveries. Used transparently, such inductive iteration has contributed to causal knowledge in comparative-historical analysis, analytic narratives, and statistical approaches. Encouraging such practices across methodologies not only adds to the toolbox of comparative analysis but also casts light on how much existing work often lacks transparency. Because successful hypothesis testing facilitates publication, yet as registration schemes and mandatory replication do not exist, abusive practices such as data mining and selective reporting find easy cover behind the language of deductive proceduralism. Productive digressions from the deductive paradigm, such as inductive iteration, should not have the stigma associated with such impropriety.

Multi-Method Research

Wawro, Gregory J., and Ira Katznelson. 2014. "Designing Historical Social Scientific Inquiry: How Parameter Heterogeneity Can Bridge the Methodological Divide between Quantitative and Qualitative Approaches." *American Journal of Political Science* vol. 58, no. 2: 526–546. DOI: 10.1111/ajps.12041

Seeking to advance historical studies of political institutions and behavior, we argue for an expansion of the standard methodological toolkit with a set of innovative approaches that privilege parameter heterogeneity to capture nuances missed by more commonly used approaches. We address critiques by prominent historians and historically oriented political scientists who have underscored the shortcomings of mainstream quantitative approaches for studying the past. They are concerned that the statistical models ordinarily employed by political scientists are inadequate for addressing temporality, periodicity, specificity, and context—issues that are central to good historical analysis. The innovations that we advocate are particularly well suited for incorporating these issues in empirical models, which we demonstrate with replications of extant research that focuses on locating structural breaks relating to realignments and split-party Senate delegations and on the temporal evolution in congressional roll-call behavior connected to labor policy during the New Deal and Fair Deal.

Other Advances in QMMR

Germano, Roy. 2014. "Analytic Filmmaking: A New Approach to Research and Publication in the Social Sciences." *Perspectives on Politics* vol. 12, no. 3: 663–676. DOI: doi:10.1017/S1537592714001649

New digital video technologies are transforming how people everywhere document, publish, and consume information. As knowledge production becomes increasingly oriented towards digital/visual modes of expression, scholars will need new approaches for conducting and publishing research. The purpose of this article is to advance a systematic approach to scholarship called analytic filmmaking. I argue that when filming and editing are guided by rigorous social scientific standards, digital video can be a compelling medium for illustrating causal processes, communicating theory-driven explanations, and presenting new empirical findings. I furthermore argue that analytic films offer policymakers and the public an effective way to glean insights from and engage with scholarly research. Throughout the article I draw on examples from my work to demonstrate the principles of analytic filmmaking in practice and to point out how analytic films complement written scholarship.

Qualitative Comparative Analysis

Krogslund, Chris, Donghyun Danny Choi, and Mathias Poertner. 2015. "Fuzzy Sets on Shaky Ground: Parameter Sensitivity and Confirmation Bias in fsQCA." *Political Analysis* vol. 23, no. 1: 21–41. DOI: 10.1093/pan/mpu016

Scholars have increasingly turned to fuzzy set Qualitative Comparative Analysis (fsQCA) to conduct small- and medium-N studies, arguing that it combines the most desired elements of variable-oriented and case-oriented research. This article demonstrates, however, that fsQCA is an extraordinarily sensitive method whose results are worryingly susceptible to minor parametric and model specification changes. We make two specific claims. First, the causal conditions identified by fsQCA as being sufficient for an outcome to occur are highly contingent upon the values of several key parameters selected by the user. Second, fsQCA results are subject to marked confirmation bias. Given its tendency toward finding complex connections between variables, the method is highly likely to identify as sufficient for an outcome causal combinations containing even randomly generated variables. To support these arguments, we replicate three articles utilizing fsQCA and conduct sensitivity analyses and Monte Carlo simulations to assess the impact of small changes in parameter values and the method's built-in confirmation bias on the overall conclusions about sufficient conditions.

Marx, Axel, Benoît Rihoux, and Charles Ragin. 2014. "The Origins, Development, and Application of Qualitative Comparative Analysis: The First 25 years." *European Political Science Review* vol. 6, no. 1: 115–142. DOI: doi:10.1017/S1755773912000318

A quarter century ago, in 1987, Charles C. Ragin published The Comparative Method, introducing a new method to the social sciences called Qualitative Comparative Analysis (QCA). QCA is a comparative case-oriented research approach and collection of techniques based on set theory and Boolean algebra, which aims to combine some of the strengths of qualitative and quantitative research methods. Since its launch in 1987, QCA has been applied extensively in the social sciences. This review essay first sketches the origins of the ideas behind QCA. Next, the main features of the method, as presented in The Comparative Method, are introduced. A third part focuses on the early applications. A fourth part presents early criticisms and subsequent innovations. A fifth part then focuses on an era of further expansion in political science and presents some of the main applications in the discipline. In doing so, this paper seeks to provide insights and references into the origin and development of QCA, a non-technical introduction to its main features, the path travelled so far, and the diversification of applications.

Questionnaire Design

Durrant, Gabriele B., and Julia D'Arrigo. 2014. "Doorstep Interactions and Interviewer Effects on the Process Leading to Cooperation or Refusal." *Sociological Methods & Research* vol. 43, no. 3: 490–518. DOI: 10.1177/0049124114521148

This article presents an analysis of interviewer effects on the process leading to cooperation or refusal in face-to-face surveys. The focus is on the interaction between the householder and the interviewer on the doorstep, including initial reactions from the householder, and interviewer characteristics, behaviors, and skills. In contrast to most previous research on interviewer effects, which analyzed final response behavior, the focus here is on the analysis of the process that leads to cooperation or refusal. Multilevel multinomial discrete-time event history modeling is used to examine jointly the different outcomes at each call, taking account of the influence of interviewer characteristics, call histories, and sample member characteristics. The study benefits from a rich data set comprising call record data (paradata) from several face-to-face surveys linked to interviewer observations, detailed interviewer information, and census records. The models have implications for survey practice and may be used in responsive survey designs to inform effective interviewer calling strategies.

Guess, Andrew M. 2015. "Measure for Measure: An Experi-
    mental Test of Online Political Media Exposure." *Political
    Analysis* vol. 23, no. 1: 59–75. DOI: 10.1093/pan/mpu010

Self-reported measures of media exposure are plagued with
error and questions about validity. Since they are essential to
studying media effects, a substantial literature has explored
the shortcomings of these measures, tested proxies, and pro-
posed refinements. But lacking an objective baseline, such
investigations can only make relative comparisons. By focus-
ing specifically on recent Internet activity stored by Web
browsers, this article's methodology captures individuals' ac-
tual consumption of political media. Using experiments em-
bedded within an online survey, I test three different measures
of media exposure and compare them to the actual exposure. I
find that open-ended survey prompts reduce overreporting
and generate an accurate picture of the overall audience for
online news. I also show that they predict news recall at least
as well as general knowledge. Together, these results demon-
strate that some ways of asking questions about media use are
better than others. I conclude with a discussion of survey-
based exposure measures for online political information and
the applicability of this article's direct method of exposure mea-
surement for future studies.

Lenzner, Timo. 2014. "Are Readability Formulas Valid Tools for
    Assessing Survey Question Difficulty?" *Sociological Meth-
    ods & Research* vol. 43, no. 4: 677–698. DOI: 10.1177/
    0049124113513436

Readability formulas, such as the Flesch Reading Ease for-
mula, the Flesch–Kincaid Grade Level Index, the Gunning Fog
Index, and the Dale–Chall formula are often considered to be
objective measures of language complexity. Not surprisingly,
survey researchers have frequently used readability scores as
indicators of question difficulty and it has been repeatedly
suggested that the formulas be applied during the question-
naire design phase, to identify problematic items and to assist
survey designers in revising flawed questions. At the same
time, the formulas have faced severe criticism among reading
researchers, particularly because they are predominantly based
on only two variables (word length/frequency and sentence
length) that may not be appropriate predictors of language
difficulty. The present study examines whether the four read-
ability formulas named above correctly identify problematic
survey questions. Readability scores were calculated for 71
question pairs, each of which included a problematic (e.g.,
syntactically complex, vague, etc.) and an improved version of
the question. The question pairs came from two sources: (1)
existing literature on questionnaire design and (2) the Q-BANK
database. The analyses revealed that the readability formulas
often favored the problematic over the improved version. On
average, the success rate of the formulas in identifying the
difficult questions was below 50 percent and agreement be-
tween the various formulas varied considerably. Reasons for
this poor performance, as well as implications for the use of
readability formulas during questionnaire design and testing,
are discussed.

Research Transparency

Symposium: "Research Transparency in Security Studies"

Bennett, Andrew, Colin Elman, and John M. Owen. 2014. "Se-
    curity Studies, Security Studies, and Recent Developments
    in Qualitative and Multi-Method Research." *Security Stud-
    ies* vol. 23, no. 4: 657–662. DOI: 10.1080/09636412.2014.970832

Research traditions are essential to social science. While indi-
viduals make findings, scholarly communities make progress.
When researchers use common methods and shared data to
answer mutual questions, the whole is very much more than
the sum of the parts. Notwithstanding these indispensable
synergies, however, the very stability that makes meaningful
intersubjective discourse possible can also cause scholars to
focus inward on their own tradition and miss opportunities
arising in other subfields and disciplines. Deliberate engage-
ment between otherwise distinct networks can help overcome
this tendency and allow scholars to notice useful develop-
ments occurring in other strands of social science. It was with
this possibility in mind that we, the Forum editors, decided to
convene a workshop to connect two different and only par-
tially overlapping networks: the qualitative strand of the secu-
rity subfield and scholars associated with the qualitative and
multi-method research project.*

Moravcsik, Andrew. 2014. "Trust, but Verify: The Transpar-
    ency Revolution and Qualitative International Relations."
    Security Studies vol. 23, no. 4: 663–688. DOI: 10.1080/
    09636412.2014.970846

Qualitative analysis is the most important empirical method in
the field of international relations (IR). More than 70 percent of
all IR scholars conduct primarily qualitative research (includ-
ing narrative case studies, traditional history, small-n compari-
son, counterfactual analysis, process-tracing, analytic narra-
tive, ethnography and thick description, discourse), compared
to only 20 percent whose work is primarily quantitative. Total
use is even more pervasive with more than 85 percent of IR
scholars conducting some qualitative analysis. Qualitative
analysis is also unmatched in its flexibility and applicability: a
textual record exists for almost every major international event
in modern world history. Qualitative research also delivers im-
pressive explanatory insight, rigor, and reliability. Of the twenty
scholars judged by their colleagues to have 'produced the
best work in the field of IR in the past 20 years,' seventeen
conduct almost exclusively qualitative research.*

Saunders, Elizabeth N. 2014. "Transparency without Tears: A Pragmatic Approach to Transparent Security Studies Research." *Security Studies* vol. 23, no. 4: 689–698. DOI: 10.1080/09636412.2014.970405

Research transparency is an idea that is easy to love in principle. If one accepts the logic behind emerging transparency and replication standards in quantitative search, it is hard not to agree that such standards should also govern qualitative research. Yet the challenges to transparency in qualitative research, particularly on security studies topics, are formidable. This article argues that there are significant individual and collective benefits to making qualitative security studies research more transparent but that reaping these benefits requires minimizing the real and expected costs borne by individual scholars. I focus on how scholars can meet emerging standards for transparency without incurring prohibitive costs in time or resources, and important consideration if transparency is to become a norm in qualitative security studies. In short, it is possible to achieve transparency without tears, but only if perfection is not the enemy of the good.*

Kapiszewski, Diana, and Dessislava Kirilova. 2014. "Transparency in Qualitative Security Studies Research: Standards, Benefits, and Challenges." *Security Studies* vol. 23, no. 4: 699–707. DOI: 10.1080/09636412.2014.970408

Discussion about greater openness in the policymaking and academic communities is emerging all around us. In February 2013, for example, the White House issued a broad statement calling on federal agencies to submit concrete proposals for 'increasing access to the results of federally funded scientific research.' The Digital Accountability and Transparency Act passed the US House of Representatives in 18 November 2013 (it has not yet been voted on in the Senate). In academia, multiple questions are arising about how to preserve and make accessible to the 'Deluge of (digital) data' scientific research produces and how to make research more transparent. For instance, on 13-14 June 2013, a meeting to address 'Data Citation and Research Transparency Standards for the Social Sciences' was convened by the Inter-university Consortium for Political and Social Research (ICPSR) and attended by opinion leaders from across the social science disciplines. In November 2014, ICPSR hosted 'Integrating Domain Repositories into the National Data Infrastructure,' a follow-up workshop that gathered together representatives from emerging national infrastructures for data and publications.*

Snyder, Jack. 2014. "Active Citation: In Search of Smoking Guns or Meaningful Context?" *Security Studies* vol. 23, no. 4: 708–714. DOI: 10.1080/09636412.2014.970409

Andrew Moravcsik makes a persuasive case that rigorously executed qualitative methods have distinctive and indispensable role to play in research on international relations. Qualitative case studies facilitate the tracing of causal processes, provide insight into actors' understanding of their own motives and assumptions, and establish an interpretive or systemic context that makes unified sense of discrete events. Teamed up with quantitative methods, qualitative methods can check on the presence of hypothesized causal mechanisms that might be difficult to measure in numerical shorthand.*

Symposium: "Openness in Political Science"

Lupia, Arthur, and Colin Elman. 2014. "Openness in Political Science: Data Access and Research Transparency – Introduction." *PS: Political Science & Politics* vol. 47, no. 1: 19–42. DOI:doi: 10.1017/S1049096513001716

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. 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. 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.

Elman, Colin, and Diana Kapiszewski. 2014. "Data Access and Research Transparency in the Qualitative Tradition." *PS: Political Science & Politics* vol. 47, no. 1: 43–47. DOI:doi:10.1017/S1049096513001777

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.

Moravcsik, Andrew. 2014. "Transparency: The Revolution in Qualitative Research." *PS: Political Science & Politics* vol. 47, no. 1: 48–53. DOI: doi:10.1017/S1049096513001789

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.

Scientific Realism/Critical Realism

Dy, Angela Martinez, Lee Martin, and Susan Marlow. 2014. "Developing a Critical Realist Positional Approach to Intersectionality." *Journal of Critical Realism* vol. 13, no. 5: 447–466. DOI: doi:10.1179/1476743014Z.00000000043

This article identifies philosophical tensions and limitations within contemporary intersectionality theory which, it will be argued, have hindered its ability to explain how positioning in multiple social categories can affect life chances and influence the reproduction of inequality. We draw upon critical realism to propose an augmented conceptual framework and novel methodological approach that offers the potential to move beyond these debates, so as to better enable intersectionality to provide causal explanatory accounts of the 'lived experiences' of social privilege and disadvantage.

Holland, Dominic. 2014. "Complex Realism, Applied Social Science and Postdisciplinarity." *Journal of Critical Realism* vol. 13, no. 5: 534—554. DOI: doi:10.1179/1476743014Z.00000000042

In this review essay I offer a critical assessment of the work of David Byrne, an applied social scientist who is one of the leading advocates of the use of complexity theory in the social sciences and who has drawn on the principles of critical realism in developing an ontological position of 'complex realism'. The key arguments of his latest book, Applying Social Science: The Role of Social Research in Politics, Policy and Practice constitute the frame of the review; however, since these overlap with those of his previous books, Interpreting Quantitative Data and Complexity Theory and the Social Sciences, I consider all three books together. I identify aspects of Byrne's ontological position that are in tune with the principles of original and dialectical critical realism and aspects that are not. I argue that these inconsistencies, which Byrne must resolve if he is to take his understanding of complexity further, stem from the residual influence of various forms of irrealism in his thinking.

Teaching QMMR

Elman, Colin, Diana Kapiszewski, and Dessislava Kirilova. 2015. "Learning through Research: Using Data to Train Undergraduates in Qualitative Methods." *PS: Political Science & Politics* vol. 48, no. 1: 39–43. DOI: doi:10.1017/S1049096514001577

In this brief article, we argue that undergraduate methods training acquired through coursework is a critical prerequisite for effective research and is beneficial in other ways. We consider what courses on qualitative research methods, which are rarely taught in undergraduate political science programs, might look like. We propose that instruction initially should involve specialized texts with standardized exercises that use stylized data, allowing students to focus on the methods they seek to master. Later in the sequence, research questions can be brought to the fore, and students can undertake increasingly complex research tasks using more authentic data. To be clear, students following the path we suggest are still learning methods by using them. However, they are beginning to do so by executing research tasks in a more controlled context. Teaching methods in this way provides students with a suite of techniques they can use to effectively and meaningfully engage in their own or a faculty member's research. We also raise some challenges to using qualitative data to teach methods, and conclude by reprising our argument.

The Journal Scan for this issue encompassed the following journals: *American Journal of Political Science, American Political Science Review, Annals of the American Academy of Political and Social Science, Annual Review of Political Science, British Journal of Political Science, British Journal of Politics & International Relations, Comparative Political Studies, Comparative Politics, Comparative Studies in Society and History, European Journal of Political Research, European Political Science Review, Foucault Studies, Governance: An International Journal of Policy Administration and Institutions, History of the Human Sciences, International Organization, International Security, International Studies Quarterly, Journal of Conflict Resolution, Journal of Critical Realism, Journal of European Public Policy, Journal of Experimental Political Science, Journal of Political Philosophy, Journal of Political Power, Journal of Politics, Journal of Women Politics & Policy, Millennium: Journal of International Studies, New Political Science, Party Politics, Perspectives On Politics, Philosophy and Social Criticism, Policy and Politics, Political Analysis, Political Research Quarterly, Political Studies, Politics & Gender, Politics & Society, PS: Political Science & Politics, Public Administration, Regulation & Governance, Review of International Political Economy, Security Studies, Signs: Journal of Women in Culture and Society, Social Research, Social Science Quarterly, Socio-Economic Review, Sociological Research and Methods, Studies in American Political Development, Studies in Comparative International Development, World Politics.*

Qualitative and Multi-Method Research
Department of Political Science
140 Science Drive (Gross Hall), 2nd Floor
Duke University, Box 90204
Durham, NC 27708