A GUIDE TO ACTIVE CITATION
Version 1.8 for Pilot Projects
(November 2013)

Andrew Moravcsik (Princeton University)[1]
Colin Elman (Syracuse University)
Diana Kapiszewski (Georgetown University)

I. Overview

This memorandum offers practical guidance on how to make the presentation and use of qualitative data in published scholarship more transparent. This is an important goal, whether the data are employed for description, narrative, interpretation, or causal inference.[2] The document suggests strategies for meeting evolving disciplinary standards for data access and research transparency. The aim is to promote openness while retaining the distinctive elements of the qualitative research tradition, and avoiding the inappropriate homogenization of social science practices.[3]

This memorandum focuses on one approach to improving data access and research transparency in published qualitative work: *active citation*.[4] This method envisages that claims and data sources that are either central to an author's argument, or controversial and contestable within the literature the author is addressing, will be backed by "rigorous, annotated citations hyperlinked to the sources themselves."[5] Active citation is thus a technologically-enabled version of traditional footnoting, citation, and appendix practices. It does not require qualitative political scientists to change fundamentally how or why they conduct research. Rather it provides a way for them to explain more clearly and systematically what they do and why they do it.

This version of *A Guide to Active Citation* is provisional. It is specifically tailored for use by scholars preparing "pilot projects" demonstrating active citation for the Qualitative Data Repository (QDR) at Syracuse University. It is a work in progress, and some gaps remain. We hope that the authors of the pilot projects will help us to develop and improve the process of active citation further, as well as to sharpen this description and guidance for future users. Hence, as you proceed with activation, it will be very helpful if you could let us know where you think the method might be carried out a bit differently, and where the instructions are vague, confusing, or lack necessary information. We hope developing these pilot projects will be an intellectually engaging task, and we are very grateful for your assistance and patience.

---

[1] Active citation was first proposed by Moravcsik (2010); see also Moravcsik (2012a). We gratefully acknowledge helpful comments on earlier versions of this document from Louise Corti (Associate Director UK Data Service) and Dessislava Kirilova.

[2] A sister document, "A Guide to Sharing Qualitative Data," supplements this document. It offers a brief set of guidelines for scholars who will be depositing data in the Qualitative Data Repository.

[3] On the issue of homogenization, see King, Keohane, and Verba (1994) as well as Goertz and Mahoney (2012).

[4] An alternative approach is "enhanced publication;" see http://www.openaire.eu/en/component/content/article/76-highlights/344-a-short-introduction-to-enhanced-publications.

[5] Moravcsik (2010), 31.

II. Background: Emerging Standards for Transparency

Norms across the social sciences are evolving to encourage greater access to data and more transparency in research practices. In political science, the American Political Science Association (APSA) is at the forefront of these changes. In October 2012, the APSA Council approved an amendment (proposed by the Committee on Professional Ethics, Rights, and Freedoms) to APSA's *Guide to Professional Ethics in Political Science* outlining revised principles for how scholars should present their research and the evidence on which it is based.[6]

The revised *Guide* specifies that researchers "have an ethical obligation to facilitate the evaluation of their evidence-based knowledge claims through data access, production transparency, and analytic transparency so that their work can be tested or replicated." These three key terms are defined as follows:

> ***Data Access:*** Researchers making evidence-based knowledge claims should reference the data they used to make those claims. If these are data they themselves generated or collected, researchers should provide access to those data or explain why they cannot.
>
> ***Production Transparency:*** Researchers providing access to data they themselves generated or collected should offer a full account of the procedures used to collect or generate the data.
>
> ***Analytic Transparency:*** Researchers making evidence-based knowledge claims should provide a full account of how they draw their analytic conclusions from the data, i.e., clearly explicate the links connecting data to conclusions.

While meeting basic standards of data access, production transparency, and analytic transparency is now incumbent on all political scientists, there is broad agreement that these general principles should be conceived and realized differently in the quantitative and qualitative research traditions. The rest of this document focuses on specific challenges and practices in the qualitative tradition.

III. How Does Active Citation Further Data Access and Research Transparency?

The qualitative research tradition would be well served by augmenting existing citation practices with a more transparent approach to the presentation and analysis of data. Active citation meets this need for greater transparency by making it easier for authors to show, and readers to see, the evidence on which authors rely and how they rely on it. This increases political scientists' ability to illustrate the precise links between evidence and arguments, and thereby to demonstrate the descriptive and explanatory richness, interpretive depth, and rigor and validity of their work. The aim of the practices described here is to make what is already good better, not to call into question existing research, or to change the way scholars write books and articles.

[6] APSA (2012)

The active citation standard seeks to encourage and extend scholarly "best practices" across the discipline.[7] Traditional practices of qualitative analysis in all social sciences and humanities—most rigorously in fields such as law, history, policy analysis, education, and classics—link evidence-based claims to the data sources that support them via citations.[8] Active citation builds upon this traditional technique. Where the evidence-based claims, and/or the data sources in citations, are critical—central to an author's overall argument, or controversial and contestable within the research community to which he or she is speaking—the author using active citation is obliged to provide supplemental information in a "Transparency Appendix" (TRAX).

TRAXs begin with an overview discussing the overall trajectory of the research, the data collection and selection procedures that were employed, and how the analysis attends to the rules of inference or interpretation that underlie the qualitative methods employed. Thereafter, each "citation entry" in the TRAX (hyperlinked from a particular citation in the piece of scholarship being activated) includes the full text of the citation, an optional "citation annotation," and eight elements (pieces of information) *for each source* referenced in the citation: (a) identification label; (b) source excerpt; (c) source annotation explaining the connection between the source and the textual claim it is supporting; (d) full bibliographic reference and additional location information for source; (e) electronic link to source material available online (optional); (f) source file (optional); and (g) information concerning shareability of provided source; (h) information concerning non-provision of source. In sum, for any reader of a scholarly work, the *prima facie* basis for a citation—basic source material and an explanation of its link to the argument—will be just one click (or, in hard copy, a few page turns) away from a scholarly citation.

Active citation constitutes a very promising means of satisfying disciplinary standards of data access and research transparency, while retaining and enhancing the richness and nuance of qualitative scholarship. Logistically, active citation provides an efficient and immediate form of electronic (or hard copy) data access for qualitative sources, without disturbing existing formats or page limits of scholarly work. Each citation can be annotated as much as is appropriate: analytic techniques, aggregation of findings, and empirical assessment can all be discussed, directly enhancing analytic transparency.

---

[7] In doing so, active citation borrows existing "best practices" in scholarly fields such as legal academia, classics, policy analysis, historiography, and natural sciences, where rigorous attention to qualitative data and data analysis has long been a hallmark. Most electronic law reviews employ annotated, hyperlinked citations to legal materials. Citations to original texts in classics journals are customarily hyperlinked through to sources. Citations to published research in natural science and medical journals are handled analogously. Some history journals have moved toward multi-media active citation, not just for text, but for scans of documents, visual and recorded material, and other items. As a result of controversy about the sourcing of their analyses prior to the Iraq War and other high-profile decisions, US intelligence agencies have also moved to active citation: even high-level analytic documents are now footnoted electronically all the way down to individual sources. Natural science journals make near universal use of appendices, often dozens or hundreds of pages long, to supply detailed data and information on methods.

[8] Throughout this *Guide*, we use "citation" as the generic term for any in-text bibliographic reference to a source, or any footnote or endnote that supplements the main text of a published piece of scholarship. Using the generic term gives us the flexibility to extend active citation to new formats that are likely to arise with electronic publishing. For example, in time, "footnotes" and "endnotes" may no longer be the appropriate terms to refer to places where information that supplements a main text is offered.

To be sure, not every reader will plumb the depths of every active citation. Many scholars will be content with the main text of a published work, and perhaps a few citations. For them, the experience of reading a scholarly work will remain essentially unchanged. Activated research will be of most direct benefit to scholars seeking to appreciate the full richness of an author's descriptive and interpretive claims. It will be particularly helpful for scholars investigating whether the information contained in cited sources supports an author's evidence-based claims. Such criticism is exceedingly rare in qualitative political science, in part because the transaction costs of such evaluation are so high; active citation would reduce these costs. Of course, it would also provide a decisive resource for authors, permitting them to display critical evidence in favor of their claims. Yet no matter how many scholars actually click open active footnotes, or use them in these ways, the emerging norm of active citation will have the collateral benefit of encouraging authors to be more careful and precise when making and supporting evidence-based claims, thus augmenting the rigor and value of qualitative political science research.

We now explain specifically how active citation furthers the three core openness goals set forth in APSA's *Guide to Professional Ethics in Political Science:* data access, production transparency, and analytic transparency.

A. Active Citation and Data Access

Active citation offers an effective and efficient means for qualitative social scientists to meet basic data access standards by linking specific arguments in the main text of a publication directly to the supporting data via hyperlinked citations. This promises to open up serious engagement in qualitative debates to a much wider body of scholars, not just a few scholars who already "know the sources." Three characteristics of active citation account for its unique effectiveness in doing so.

First, *by requiring precise references, active citation allows scholars to identify and locate the sources which underlie their claims.* The best traditional citations carefully identify supporting sources, and where the relevant evidence is located in those sources. However, all too often, authors provide incomplete or vague information about particular sources and/or the location of evidence within them, making it difficult for other scholars to determine whether the cited evidence supports the author's claims. Even such basic components as page numbers are often missing, as well as what aspect of the cited material supports the claim in the text. Some "scientific" citation methods do not permit scholars to display sufficient information to locate various types of primary and archival documents, for instance. Active citations address this problem by assuring that sufficient information is provided so readers can locate cited sources and find the evidence to which an author is pointing to substantiate his or her claims. This is possible while preserving existing citation formats.

Second, *active citation places data directly before the reader.* Even where traditional citations precisely identify the location of a source and of the relevant evidence it contains, readers must fetch the materials for themselves. There is no tradition of (and no supporting infrastructure for) providing access to cited sources. Thus while in theory precise citation makes it possible for readers to access materials and evaluate to what degree they support an author's claims, in

practice the transaction costs are often so high that empirical claims made in published political science research are rarely checked. Facilitating readers' access to the underlying sources themselves is particularly important in cases where qualitative scholars generate their own data, for example via interviews, archival research, or participant observation, rather than relying on existing datasets.[9] Active citation requires scholars to provide excerpts from the sources they cite (and ideally the sources themselves), thereby dramatically reducing the costs of evaluating the extent to which an author's evidence substantiates his or her claims. With active citation, readers have before them the evidence required to assess the *prima facie* empirical plausibility of any claim and often to place it in a broader interpretive context. Active citation can thus facilitate the evaluation and replication of research results in qualitative political science by decreasing the transaction costs of acquiring the data that underlie them.

Third, under some circumstances, *the active citation approach facilitates re-use of data for secondary analysis.* In quantitative political science, scholars often reuse data collected and deployed by other scholars in order to extend an analysis or launch new research projects. The traditional availability of datasets in electronic form makes it relatively easy for quantitative scholars to do this. Secondary analysis remains much rarer in qualitative traditions, in large part because the evidence underlying existing studies is more costly and difficult to access. Every scholar begins, as it were, from scratch. In some instances, active citation can empower reuse, because an important subset of the sources cited in a publication (or at least an excerpt therefrom) can be made available via the TRAX. Potential secondary users of evidence need not find each source individually, but can immediately examine, download or otherwise manipulate evidence cited in an article or chapter for whatever scholarly purpose they wish. Obviously, the feasibility and utility of secondary analysis of the data that are shared via active citation will depend on how much information was made available, and to what degree those data constitute a complete and coherent set.

B. Active Citation and Production Transparency

The primary purpose of active citation is to clarify connections between claims that are supported by in-text references, footnotes, or endnotes, and their underlying data sources. In order for the research design, descriptive and causal inferences, and interpretive choices to be clear to other scholars, however, they must understand how and why those *particular* data (and not others) were generated or collected, how and why they were used in the analysis, and why they were cited and others were not. The standard of *production transparency* requires authors to explain the origin and generation of data they cite in their publications.

This is important in qualitative analysis, where the potential for particular biases in information collection and data generation (such as sampling bias) often looms large. One common form of bias results from "cherry-picking" evidence (i.e., over-representing evidence confirming a favored hypothesis). This can arise purposefully or inadvertently; and it can arise whether a

[9] This is true even with published secondary sources, which may be difficult to find or be written in foreign languages. Primary source material may have been collected from multiple remote archives, for which a cumbersome process is required to attain the requisite permissions and similar linguistic barriers arise.

scholar employs primary or secondary, oral or written, textual or visual sources.[10] Such biases are generally difficult to spot because, in current practice, the absence of reference to evidence favoring other explanations is an error of omission invisible to all but the most expert scholar who already knows the literature or data. While one can never eliminate entirely the "cherry picking" risk, rendering data-generation procedures more transparent should be a goal. Doing so encourages authors to be more self-conscious about how they collect information and about interpretations of subjective experiences, renders research results more legitimate and tractable, and facilitates scholarly debate and criticism. Of course, some aspects of production transparency are generally addressed in the text of books and articles. However, word limits and other constraints, as well as the lack of a general expectation that scholars be fully transparent, means that scholars seldom fully explain their data-generation procedures.

Providing a TRAX and engaging in active citation can increase production transparency by supplementing (rather than replacing or repeating) what is said in the text. Scholars can offer additional detail on the project's empirical base in the overview section of the TRAX, providing a holistic, synthetic discussion of the context of data collection, and elaborating the procedures used to access, collect, generate, capture, prepare, and select data for inclusion as evidence.

C. Active Citation and Analytic Transparency

Active citation's greatest strength is its ability to facilitate *analytic transparency*. It does so by rendering patent the links between an evidence-based interpretation, or descriptive or causal inference, in a text, and the particular data on which it is based. In doing so, active citation builds on traditional citation practices, thereby preserving unique epistemic characteristics of qualitative research. Active citation improves analytic transparency in three ways.

First, *active citation links individual pieces of evidence to aspects of an inferential or interpretive analytic process.* Perhaps the most ubiquitous qualitative method is process tracing, where scholars identify "causal process observations," defined as "an insight or a piece of data that provides information about context or mechanism and contributes…leverage in causal inference."[11] Such data are not "variable scores to be assembled into a rectangular dataset," but rather comprise a series of observations linked to specific stages, aspects or components of a narrative, process, or sequence, or used to test specific counterfactual hypotheses about mechanisms within cases.[12] By allowing scholars to connect particular pieces of evidence to specific stages of the causal process being discussed, the active citation format makes especially transparent the way inferences are being drawn from that connection. This method can be tailored to the characteristics of the qualitative data and the type of inference the author seeks to draw.

Second, *active citation places arguments in a broader evidentiary context.* Traditional citations link arguments to a reference or, at best, a very short quotation. By contrast, active citations link

[10] Lustick (1996).
[11] Collier, Brady and Seawright (2010) 184. See also Mahoney (2012), Waldner (2012).
[12] Goertz and Mahoney (2012) 88-90.

arguments to an excerpt from the source(s) the author is positing supports the arguments. The excerpts are of sufficient length for a reader to judge whether the source actually supports the interpretation, or descriptive or causal inference for which it is being called on as evidence in the main text. Particularly when the sources themselves are also provided, active citation offers some assurance that scholars do not take evidence out of context. More subtly, and of paramount importance from an interpretivist perspective, active citation can provide the reader more direct access to the voices of the actual subjects of analysis and a richer understanding of the specific cultural, strategic or social context from which particular social facts emerge.

Third, *active citations can be annotated, offering researchers an opportunity to "show their work."* Activated citations are generally accompanied by an annotation describing more precisely how the author used the data cited to arrive at the descriptive claim, interpretation, or causal inference discussed in the main text. They thus differ from traditional (un-annotated) citations (and, in particular, from increasingly widespread "scientific" citation styles), which simply juxtapose claims and references without explaining the precise connection between them, thus requiring that readers engage in "reverse engineering" to identify and evaluate the link being posited. As Moravcsik notes, in order to effectively evaluate evidence-based claims, readers need to know "exactly how and why the citation supports the textual point."[13] Further, active citations permit scholars to document a wider and richer range of descriptive, interpretive and causal arguments. Authors may use annotations to underscore the strength of such arguments, or to acknowledge sources of complexity, uncertainty or ambiguity. Active citations may also be used to anticipate and answer criticisms of an inference or interpretation, or to clarify methodological issues not fully explained in the text or footnotes.

IV. Active Citation: Standards and Practices

This section provides guidance on practical issues and standards for research using active citation. It discusses when scholars should engage in active citation; which citations in a particular piece of scholarship should be activated; how they should be activated; and the format of the Transparency Appendix (TRAX).

A. When Should Scholars Engage in Active Citation?

In accordance with the new APSA standards, active citation can be used to facilitate data access, production transparency, and analytic transparency in all qualitative political science research. Scholars can activate the citations in a piece of scholarship and produce a TRAX at any point (including as they are writing, when it will be easiest to do so).

The APSA guidelines provide that scholars may enjoy first-use of their data unless required to release them earlier (for example, by funders, journals, conferences, hiring committees, or publishers). That is, unless otherwise mandated, scholars have one year from the date of

[13] Moravcsik (2012) 34.

publication to provide access to the data underlying that publication, and a description of how they were generated or collected (production transparency).

Even when disclosure is not mandated by third-parties, scholars may choose to forego their first-use window. There are several good reasons not to delay sharing the data that underlie an activated piece of scholarship. From a logistical standpoint, it may be costly for publishers and editors to issue books and articles and then to revisit them a year later to add hyperlinks and sources. From an intellectual standpoint, it is often only possible to assess the *prima facie* rigor and plausibility of evidence-based claims if the data sources and arguments that underlie them are available. If they are not, journal and book reviewers, and editors, must reach a decision whether to accept work for publication without being able to evaluate its true quality. This not only foregoes opportunity for critique and improvement, but risks embarrassment if the sources are found to be inadequate once revealed.

B. Which Citations Should Be Activated?

Scholars occasionally express concern about the additional burden using active citation entails. As we noted above, in many respects active citation is simply an updating of best practices in traditional qualitative research. Accordingly, it involves only a modest increment in effort for scholars who are already systematic and careful.

One of the most difficult tasks is determining which citations need to be activated. We suggest the following standard: only those citations that are central to an author's argument, or controversial and contestable within the literature the author is addressing, need be activated. What does this mean? We propose four criteria to assess which citations should be activated in a scholarly work, and consider different types of citations that might appear in such work (see Table 1).[14] If one or more of these criteria are fulfilled, we encourage fully active citation.

(1) **Centrality of the (Evidence-based) Claim:** Citations to evidence-based claims that are crucial building blocks to an author's overall argument or thesis should be activated;

(2) **Importance of the Data Source:** Citations referencing crucial pieces of evidence underlying the author's overall argument or thesis should be activated;

(3) **Contested or Controversial Nature of the (Evidence-based) Claim:** Citations to evidence-based claims that are contested or controversial within the research community, debate, or literature to which an author is speaking or on which he or she draws, or might be questioned by members of those communities,[15] should be activated;[16]

[14] APSA's amended standards for data access and research transparency were developed for application to published research. Accordingly, the discussion in this section focuses on published books and articles. Although it is not a formal requirement, to the extent authors make evidence-based arguments in working papers and other unpublished material, we believe it is best practice to also use active citation in those formats.

[15] Requiring that authors activate citations to claims that might be controversial not just within the discipline *to* which the article is directed (e.g., political science) but within any discipline *from* which it draws material should offer some protection against cherry-picking results from other disciplines (Lustick 1996).

(4) **Contested or Controversial Nature of the Data Source:** Citations referencing data sources that are controversial within the research community, debate, or literature to which the author is speaking and on which he or she draws, or whose validity or evidentiary value might be questioned by members of those communities, should be activated.

We recognize that these criteria have some associated difficulties. For example, all involve judgments by the author: the degree to which claims or data are important or controversial is not exact, and may be specific to a particular time (i.e., what is important or accepted at once moment may be called into question later). We expect that the judgments entailed in engaging in active citation will be made in good faith, and also hope the criteria provide a framework in which active citation can be challenged by other members of the research community. We believe these criteria make it more likely that the activated citations will deliver useful transparency, while being mindful of the opportunity costs involved in active citation, and hence the pragmatic limitations on what it is reasonable to expect authors to undertake.

Of course, the rules set forth above are just one rendition of what such a set of criteria might look like. A continuum exists – from fully activating all footnotes to not activating any – and different scholars will be more comfortable with different scenarios. Scholars could, for instance, activate all footnotes that reference anything other than a brute fact, or activate any that involve descriptive or causal inference no matter how big or small. We invite those developing QDR pilot projects to determine their own logic and activate the citations they think important to activate (keeping in mind the goals of the overall transparency project). In order to help us to develop optimal standards, we ask that scholars who use strategies that deviate from the one we propose here clearly explain and justify, in the TRAX overview, the logic by which they chose citations to activate.

To further develop the "continuum" idea just mentioned, we refer again to Table 1, which distinguishes among four types of citations: (1) a traditional citation, (2) a fully active citation (with both annotation and source material); (3) an annotated traditional citation; and (4) a sourced active citation. We elaborate on the circumstances appropriate to each below.

1. The ***traditional citation*** simply contains a reference, but need include neither annotation nor source material. Traditional citations remain appropriate when none of the criteria outlined above is met. For example, the citation may simply reference alternative social science theories, provide methodological guidance, acknowledge previous work on a topic, offer general empirical background, discuss related examples the author does not intend to develop in detail, establish the general importance or policy relevance of a study, or set the mood by

[16] Consider a few examples. While the dates of President Obama's inaugurations may seem obvious, he in fact was sworn in on two different dates for each term, in part as a result of the complex constitutional consequences of Justice Roberts' having misspoken at his initial inauguration. The true population of China and other countries is the subject of considerable debate. The word count of the Declaration of Independence depends on which words one counts: signatures or not, headings or not. These ambiguities may or may not matter for any given study. A study of the constitutional history and politics of inaugurations might require source material on these points, whereas a study of inaugural speeches, for which the date is simply a peripheral fact, may not require an annotation or a source, or even a citation.

providing epigrams. By contrast, if the argument of the study in question specifically and centrally problematizes the empirical veracity of these items, or rests on specific, questionable assumptions about them, an active citation is likely to be appropriate.

Table 1: Four Types of Citations

		ANNOTATION PROVIDED	
		YES	**NO**
SOURCE PROVIDED (AT LEAST AN EXCERPT)	**YES**	Fully Active Citation (2)	Sourced Active Citation (4)
	NO	Annotated Traditional Citation (3)	Traditional Citation (1)

2. The ***fully active citation*** is recommended as the standard default format for any claim or data meeting one or more of the centrality/controversy criteria mentioned above. The citation (in the format of the publication) is hyper-linked to a citation entry in the TRAX containing several items including an excerpt from the source at least 50-100 words long (presumptively, in the case of textual sources), and an annotation. This requirement holds whether the text is a primary document, a redaction or summary of a document, a transcript of an interview or focus group, an oral history, a set of ethnographic or participant-observer field notes, a diary and other personal record, a press clipping, a secondary publication, etc. The source length requirement may be adjusted due to human subject, intellectual property, and logistical constraints—all of which trump the presumptive length requirement. It is often beneficial for the fully active citation to be accompanied by a copy of the textual or non-textual source itself (e.g., a scanned document, audio or video recording, photograph, map, poster, artwork and other visual artifact, etc.). These points are discussed in more detail below.

3. The ***annotated traditional citation*** is simply a citation (formatted as customary in the publication) containing a reference and an annotation. This type of footnote, which has long offered scholars a means of elaborating the relationship between evidence and argument, was traditionally employed in history and remains the norm in law, classics and other fields. Its use has declined recently in political science as "scientific" citation has supplanted classic citation form and overall word limits have grown tighter. An annotated traditional citation is appropriate when such citations are not permitted by a publication or space limitations restrict their use, *and* when the cited source documents an unambiguous and uncontested historical fact but the relationship between the claim in the main text and the citation requires elaboration. When a causal claim is contestable, however, and residual ambiguity exists about the content, veracity or representativeness of the cited source, a fully active citation is preferable.

4. The ***sourced active citation*** is hyper-linked to a citation entry in the TRAX including an excerpt from the source (and, ideally, the source itself), without annotation. Providing the source without an annotation may be appropriate under a variety of circumstances. The connection between the source and the claim may already be clearly stated in the existing

citation. More rarely, the meaning of the source and how it supports the evidence-based claim in the text may be unambiguous and uncontestable simply based on its juxtaposition with the text and citation. We expect this will happen infrequently because the meaning has to be apparent not just to the author who is typically familiar with a scholarly interpretation but also to her readers who may be coming to it for the first time. We anticipate that it will very often be preferable to provide a very short (e.g., one sentence) annotation to maximize the analytic transparency of the claim – that is, to fully activate the citation.

C. The TRAX

Engaging in active citation involves creating a TRAX – a standardized means to present information about data, how they were selected, and how interpretations and inferences were drawn from them. Creating a TRAX assists scholars in efficiently conforming to standards of data access, production transparency and analytic transparency. The TRAX is normally created contemporaneously with writing a paper, article or book chapter (although for pilot projects based on previous work it is being generated retroactively).

QDR is developing two software tools to help scholars creating QDR pilot projects to generate their TRAXs. First, the Active Citation Editor (ACE), which can be accessed on the QDR site and is ready for use, is designed to help scholars to activate the citations in already-published work.[17] The Live Active Citation Editor (LACE), a Microsoft Word plug-in that is still in development, aids scholars in activating citations simultaneously with writing. Both allow scholars to focus on the content of the TRAX rather than on mechanics such as formatting.

QDR prefers that scholars who are activating the citations in already-published scholarship use the ACE so that TRAXs posted to QDR will be standardized. We look forward to receiving your feedback on this tool and how it might be improved. Of course, we understand that not everyone has the same elective affinity for using new software tools; scholars who do not wish to use the ACE should contact QDR staff to discuss alternative ways to create their TRAX.

Before continuing, we wish to note that we understand that engaging in active citation retroactively (i.e., drawing on material collected for and employed in projects that have been completed) is difficult for several reasons. Authors may not have formally recorded their data-selection and data-collection procedures, or may not have retained their notes. Moreover, sampling techniques and choices, analytical and interpretive decisions, and the trajectory of research are often driven by a complex mix of intellectual and pragmatic considerations, which may be difficult for scholars to reconstruct years later. We hope authors will do their best to recall the research process.

How detailed the TRAX will be depends upon the type of research project and publication, and the researcher's ambition with regard to rigor and precision.[18] There is no single rule.

[17] An ACE User Manual will soon be available on line to assist piloteers in using the ACE to create their TRAX.
[18] For a project replicating an existing study in order to question the data analysis, or applying a set of established and consensual theories, hypotheses and inferences to new data, the focus may be largely on data collection and data

Nonetheless, all TRAXs will include two sections: an introductory "overview," and a series of "citation entries." We treat each in turn below.

(1) *Overview*

Achieving transparency requires explaining the logic of inquiry and evidentiary basis for the researcher's conclusions. Some aspects of this openness do not readily attach themselves to a particular citation, but rather are best dealt with holistically. When such data-related issues and methodological questions cannot be addressed comfortably in the main text of a publication, scholars may expand on them in the introductory overview of their TRAX. This section of the TRAX is equivalent to a methodological appendix in the quantitative or formal traditions of political science, and also found widely in other natural and social sciences. Specifically, scholars can discuss their overall research trajectory (e.g., how interpretations and hypotheses were generated and evaluated); provide additional information about the procedures used to collect/generate, prepare, and select data for inclusion in the publication; and illustrate that they followed the rules for analyzing or interpreting evidence and data that attend the analytic methodology they have employed.

a) Scholars can supplement the literature review, theory, and hypothesis formation sections of published work by further discussing their *research trajectory*. How did they select theories, and generate, evaluate, and modify hypotheses, interpretations or narratives in the course of research? By what means were observable implications derived from theories or interpretations and how did the researcher seek to observe them? Why and how were individual findings aggregated to confirm and disconfirm hypotheses, and how were these conclusions in turn aggregated into claims confirming and disconfirming broader theories? Scholars will be better able to effectively describe their research trajectory (and thus be more transparent) the more carefully they document the choices they make, the procedures and practices they use, and the sequence in which they employ them *as they carry out* their research.[19]

b) The overview supplements the discussion of *data collection* in the text and footnotes. In much of political science scholarship, readers only see the evidence that is finally cited, raising the possibility of purposeful or inadvertent "confirmation bias." While one cannot eliminate this possibility, production transparency recommends that authors describe their projects' empirical base and offer a holistic, synthetic discussion of the context of data collection, and how they were accessed, collected, generated, captured, prepared, and selected for inclusion as evidence in their publication. What is the overall field of potential evidence and what justifies decisions to select or sample particular evidence to

---

analysis. A study that adds an original theory or set of hypotheses to an existing debate about a case may focus more on theory and hypotheses formation.

[19] In the natural sciences, the evolution of research projects can be traced retrospectively through time. This is essentially the function of the experimental "lab notebook." Notebook in hand, a researcher need only distill and systematically synthesize its content in order to describe how s/he generated and evaluated interpretations, data and inferences, and arrived at conclusions. While describing the entire research process from beginning to end is not (yet) a realistic expectation for political scientists (and is not a requirement here), it should be considered an aspirational goal.

examine, to analyze in detail, and to cite?[20] What relevant data were not or could not be consulted and what is their potential effect on the analysis? What instruments and procedures were employed to collect data?[21] If interactive data-collection techniques such as interviews were used, what procedures were employed to identify and access interview respondents; what terms of confidentiality and informed consent were employed; and what interview protocols, topic guides and instructions were issued? Were data photographed, scanned, photocopied, recorded, or summarized in notes? How were data prepared for sharing? Data-collection practices, and the broader research process, are often messy, and discussions thereof in the TRAX will have both practical and methodological aspects.

c) Authors can enhance discussion of *data analysis* in the text and footnotes by offering more information about how inferences were drawn and claims were derived from data. The analytic methods that are typically employed in a qualitative research design (or a multi-method design with a significant qualitative component) can include process tracing, counterfactual analysis, pattern matching, congruence testing, small-n comparison, or *fs/*QCA and other set theory approaches. These different methods imply different approaches to (and thus different rules for) drawing descriptive or causal inferences, and each relies on a set of underlying principles or logics that justify its use and support the claim that properly employing the method facilitates drawing valid descriptive and causal inferences. While methodologists differ concerning *which* rules to follow, most agree that following a particular set provides scholars with more inferential leverage than they would otherwise have. The APSA standard likewise prescribes no epistemology or methodology; it simply requires that whatever approach is chosen, analysis be conducted as transparently as possible – i.e., that scholars clearly demonstrate that they have followed the general rules that attend the methodological approach they are using.

(2) *Citation Entries*

Recall that we use "citation" as the generic term for any in-text bibliographic reference to a source, or any endnote or footnote that supplements the main text of a published piece of

[20] Scholars might explain, for example, how they moved from the "data field" (all of the information that is potentially relevant to the question under investigation) to "consulted data" (data the author considered and consulted while conducting research, formulating the question and basic claims) to "heuristic data" (data the author considers part of the relevant sample formulating an description, narrative, interpretation, or test of particular claims) and finally to "cited data" (data actually referenced in published work). For example, if one were studying a British foreign policy decision, one might indicate that one consulted the private papers and memoirs of the Foreign Secretary, Hansard (the House of Commons official report), and interviewed several former and current government officials. One would also justify why the particular private papers, reports, and interview respondents are relevant to one's question, why they can be seen as an unbiased sample of a larger body of evidence, consider the biases that might arise from not examining the views of other ministers, lower officials, parliamentarians, social actors, and other actors, explain how one interpreted or coded the evidence from these sources, and account for the selection of particular evidence for citation.

[21] Scholars who wish to provide hyperlinks to research materials (for instance, survey questionnaires or interview protocols) should submit those materials to QDR via Dropbox; please see the Dropbox procedures described in the penultimate section of this document.

scholarship. Each citation being activated has one corresponding citation entry in the TRAX, to which it is hyperlinked. Typically, a citation contains at least one bibliographic reference to a source. However, because citations may list multiple sources, any TRAX citation entry may also include information about multiple sources. Citation entries appear in the TRAX in the same order in which the citations being activated appear in the published work.

Each citation entry begins with the **relevant citation**. For footnotes and endnotes, that will be the full note. For in-text references, that will be the portion of the main text (usually the immediately preceding sentence or two) in which the references are invoked.

If they wish, scholars may also include a "**Citation Annotation**." A citation annotation (as opposed to a "source annotation," described below) provides supplementary information that is related to more than just a single source. A citation annotation might serve a range of functions, including discussing how the various sources referenced in a citation work together to collectively support a claim in the text, or describing the relationship that the sources referenced in that citation *and another* citation have to a textual claim.

Each TRAX citation entry then includes, *for each source* referenced in the relevant citation, eight elements in the following order: (a) identification label; (b) source excerpt; (c) source annotation; (d) full bibliographic reference and additional location information for source; (e) electronic link to source material available on-line (optional); (f) source file (optional); (g) information concerning shareability of provided source; (h) information concerning non-provision of source.

a) Identification Label. This label indicates the place in the article or book chapter where the citation is located. For scholars using the ACE, the identification label is generated automatically.[22]

- *For publications with numbered citations:* For books, the identification label includes the chapter number, citation number, and a letter for the source. For example, if several sources are referenced in Chapter 7, Citation 24, the first source would be labeled 7-24-a, second source 7-24-b, and so on. For articles, the identification label includes the citation number, and a letter for the source. For example, in Citation 24, the first source referenced would be 24-a, second source would be 24-b, and so on.

- *For publications with non-numbered in-text references:* For books, the identification label includes the chapter, page number, and a letter for the source. For example, in Chapter 7, page 300, the first source mentioned would be ITR-7-300-a, the second source would be ITR-7-300-b, etc. For articles, the identification label includes the citation number, and a letter for the source. For example, the first source referenced on page 212 would be ITR-212-a, the second source would be ITR-212-b, and so on.

b) Source Excerpt. For the purposes of QDR pilot projects, brief excerpts are transcribed from textual sources. This makes the TRAX comprehensive and self-sufficient (i.e., inclusive of excerpts from and annotations to all sources); renders the source material

[22] We are still developing a proper format for activating the citations in electronic books without page numbers.

within it easily readable and fully searchable for scholars interested in locating, examining, reusing or downloading source material; and eliminates problems of outdated, inoperative, unstable external links, as well as unreadable external documents, which bedevil efforts to hyperlink footnotes in other fields.[23]

How long should the source excerpt be? For all sources, textual and non-textual, four general standards govern the proper amount of the source to be excerpted:

(i) The excerpt should clarify the nature of the evidence and provide sufficient context to interpret that evidence properly and with nuance.
(ii) The excerpt should not violate intellectual property or human subject protections.
(iii) The length of the excerpt should not impose a logistically prohibitive burden.
(iv) Within these constraints, scholars should provide as much evidence as possible.

We recommend a *presumptive minimum length of 50-100 words for textual sources*. This length is often minimally sufficient to convey the general meaning and context, lies within current scholarly practices with regard to "fair use" of most types of intellectual property, and imposes a manageable burden with regard to transcribing and anonymizing sources for human subject purposes. Scholars enjoy discretion in interpreting this presumptive norm, since circumstances will surely arise where it is appropriate to fall short or not cite anything at all.[24] When this occurs, the author should explain why the excerpt diverges from the presumptive standard. Exceeding the standard is encouraged and (keeping in mind human subject and intellectual property requirements and the relevance of the material) requires no explanation.

Are there alternatives to a quoted excerpt? In some cases, an author may be unable to provide a full verbatim text of the presumptive length. For example, an author may not possess a source in verbatim form (e.g., no interview notes were taken), it may be lost, or may not be shareable due to human subjects or copyright concerns. In such cases, the author should offer an explanation, and provide as much material as is appropriate and feasible, for example, research notes, or a substantial redaction or summary of the information on which the claim in the text is based.

What about sources in foreign languages? If the source is in a foreign language, it must be presented both in the foreign language and in translation in the language of the book or article. The name, date and reference information (if any) of the translation should be given. Access to the original language enhances research transparency. We recognize that this

[23] At the present time, excerpting source material in the TRAX is the best way to assure an adequate level of data access and analytic transparency. With technology and broader data practices moving at a rapid pace, however, it is possible that this will change in the future for some sources. A reliance on external links will be possible once the following obtain: permanent and stable location; swift means to examine, search and download single and multiple annotations and sources; uniformity of treatment and format between hyperlinked and traditional sources; translation and legibility; and a means for authors to highlight or excerpt critical passages or sections in a larger document or website. Until such a software solution exists, excerpting text in the TRAX remains the best way to implement active citation.

[24] For example, certain types of intellectual property, such as poems or art works, are often subject to more stringent "fair use" intellectual property restrictions.

imposes a logistical limit on length, and thus the translation requirement applies only to the presumptive minimum source text; the author may apply it to additional text, scans, or other materials that are provided at his or her discretion.

What about non-textual sources? Active citation guidelines are still being developed for non-textual sources such as visual representations, music, etc. As such, there is not yet any presumptive minimum standard analogous to the 50-100 word rule.

What about permissions and intellectual property issues? We expect that for most textual sources, the minimum requirement will fall under "fair use" for non-profit purposes and will not require special permission. However, we are still formulating firm guidelines on this issue. For the purpose of pilot projects, QDR will seek to obtain any necessary permissions from authors, editors and publishers to reproduce the text, references, and sources. Where this is not possible, we will advise use of as much material as is permitted by fair use.

c) Source Annotation: Source Annotations are mandatory when the relationship between the cited source and the claim in the text is not patent from the text and the existing citation (in the case of endnotes and footnotes). When this is the case, a source annotation specifies how the cited evidence relates to and supports the inference or interpretation drawn in the main text and/or citation. Annotations may be provided on a voluntary basis if, for example, an author would like to amplify a certain point or position his claim vis-a-vis that of another scholar. Annotations are more likely to be needed in TRAXs accompanying articles published in journals that mandate in-text references and discourage footnotes and endnotes; more generally, they are more likely to be needed in articles than books (although they may increasingly be necessary in the latter as publishers implement more stringent length limitations).

What exactly is in the annotation? Annotations describe the analytic path between cited data and evidence-based claims in the text. The precise content of an annotation depends on the type of claim the author seeks to substantiate; the particular principles, logics, or rules of the analytic methods he or she is employing; the type of data; and the extent of ambiguity or complexity in the evidence. Annotations might highlight both confirming and disconfirming evidence, and/or point out subtleties, nuances, contradictions and ambiguities in the evidence. If the claims in the text are contested, the annotation may discuss evidence that *refutes* potential criticisms, perhaps referencing multiple and competing sources not cited in the text. In the context of legacy footnotes activated subsequent to publication, annotations might be used to add information that emerged after publication.

Because this is a project in development, it is not yet possible to offer an exhaustive list of annotation types; indeed, we hope to develop a more complete list with the help of the pilot projects. Nonetheless, in general terms, we can distinguish between annotations supporting descriptive claims, and annotations supporting causal claims.

Annotations supporting *descriptive claims* supplement the text by making clear how the author is interpreting the cited data such that they are evidence of a particular descriptive assertion. Such annotations can substantiate a range of descriptive assertions, from simple

statements of fact ("a" occurred on a particular date) to more nuanced or contested descriptive or interpretive assertions concerning the place or role of a particular element or node in a narrative, or the relationship between two elements or nodes, for instance, "a" preceded "b;" "a" followed "b;" "a" happened in a certain way; "a" was unusual; "a" was important; "a" meant a certain thing; "a" was interpreted in a particular way by a community, etc.

Annotations supporting *causal claims* supplement the text by making clear how the evidence shows that the asserted explanatory relationship holds. The specific content of these annotations depends on the analytic methods an author is employing and their particular principles, logics, or rules for drawing causal inferences, and the type of data he is using. To give just three examples, these annotations may show how "a" was necessary for "b;" how "a" was sufficient for "b;" or that a causal mechanism was operating in the way the author asserts. Put differently, the author's methods determine what he must do in order to make a persuasive causal argument; the persuasiveness of his causal claims depends on the degree to which his presentation and discussion of evidence (in the text and annotations) suggests that he has done so.

How long should annotations be? Annotations need not be long, but there is no hard and fast rule. The appropriate length and detail may range from a one-sentence statement of the main point to a detailed acknowledgement of controversy, ambiguity, or complexity in the relationship between text and source. If plain reading of the text, footnote, and source material make self-evident how the source relates to the interpretation or inference in the main text, an extremely short annotation pointing to a critical passage or underscoring a relatively simple interpretation may be sufficient. Because situations in which no ambiguity exists (even in the mind of a reader without background knowledge) are rare, the recommended minimum is a very short annotation.

d) Full Bibliographic Reference and Additional Location Information for Source. The full bibliographic reference should be sufficiently precise so that readers can locate the source and find within it the exact passage of textual evidence to which an author is referring (i.e., page numbers). The full bibliographic reference may be longer and more detailed than the corresponding reference in the citation in the publication being activated. Scholars can also include additional information that may have been excised from the citation in the publication being activated, or that is otherwise necessary to allow others to locate the source and find the relevant information within it. For instance, for primary archival sources, information about the archive, collection, and number of the box in which the document was found should be noted.

e) Electronic link to Source Material (*optional*): Authors may supplement their source excerpts with links to the actual sources if the sources are available on-line. To be clear, such hyperlinks are not a *substitute for* but rather a *supplement to* the required excerpt.

f) Source File (*optional*): Where possible, authors are strongly encouraged to provide a scan of all or part of the original source (be it a primary document or secondary source; transcripts of interviews, focus groups, or oral histories; ethnographic or participant-

observer field notes; diaries and other personal records; press clippings; research notes; photographs, maps, posters and other representational work, or artwork).[25] For textual sources, while ideally the author would provide the entire source, a suggested minimum standard for partial provision might be three pages before and after the excerpted text. The author may also provide copies of non-textual materials, such as audio or audiovisual recordings, artwork, maps, diagrams and other representation material, and other items. Here again we welcome piloteers' input on the optimal amount of additional source information to provide.

g) Information Concerning Shareability of Provided Source: If the source is provided, authors should indicate whether it is shareable via QDR. If it is not, they should describe the nature of the constraint (e.g., human subjects concerns, copyright restrictions, etc.) Even if the source is under external constraint, it is still helpful for QDR to have access to copies of the relevant pages for offline, internal use (e.g., so we can check the excerpt that you provide for accuracy). Furthermore, a source that is currently not shareable may become so in the future. At that point, it will be easier to hyperlink the TRAX to the source if it is already stored in QDR.

h) Information Concerning Non-Provision of Source. When scholars provide neither a link to the underlying source nor a source file, they should indicate the constraints that prevent them from doing so – for example, human subjects agreements, copyright law, logistical circumstances, or another external constraint. For documents under copyright constraints, it would be helpful for scholars to note current copyright ownership.

Table 2: Provision and Shareability of Original Source

		ORIGINAL SOURCE IS SHAREABLE	
		YES	**NO**
ORIGINAL SOURCE PROVIDED TO QDR	**YES**	TRAX includes source excerpt and, hyperlinks to a PDF of original source	TRAX includes source excerpt; QDR uses original source internally to check accuracy of excerpt and retains source for potential future sharing
	NO	TRAX includes source excerpt and hyperlink to third party site on which original source is available.	TRAX includes source excerpt

V. Using the ACE to Create your TRAX

For each pilot project, our intention is to make available on-line through the QDR (1) the original publication (if permission has been secured from the publisher/press); (2) the TRAX; (3) any underlying source materials that are neither sensitive nor under copyright.

[25] See below for a discussion of how to organize and label these sources. This Guide's sister document, "A Guide to Sharing Qualitative Data," has additional information on types of qualitative data and how to prepare them for sharing.

As we noted above, we strongly encourage scholars working on active citation pilot projects to use the software tool that QDR has developed, the ACE (Active Citation Editor), to create their TRAX. The ACE is available on the QDR web site (here), which will go live by the end of November 2013.[26]

As soon as you are ready to use the ACE to create your TRAX, please notify QDR (email Diana Kapiszewski, dk784@georgetown.edu) so we can establish a Dropbox folder for your pilot project. Once we do so, please upload the following materials there:

1) A PDF of the article or book chapter that you are activating;[27]
2) The "overview" document that you would like to appear at the beginning of your TRAX;
3) Non-sensitive source materials (see below for naming convention);
    a. For textual sources, if the whole source cannot be provided, a suggested minimum standard for partial provision is three pages before and after the excerpted text.
    b. If you cite particular pages from a certain source more than once, you only need to provide one PDF of the source. That is, if you cite pages 32-36 of a source in one part of your work, and pages 35-36 in another part of your work, you can just post to Dropbox one PDF including pages 32-36. Likewise, if you also cited pages 57-59 from the same source somewhere else in your work, you can send us one PDF including pages 32-36 and pages 57-59.
    c. Please do not use Dropbox to transfer sensitive source materials – we will establish different practices for the transfer of such files. Please contact QDR staff for instructions about how to transfer these files.
4) If you are activating the citations in a single chapter from a monograph, a PDF of the book's bibliography.

Once we receive your materials, QDR staff will set up your project on the ACE so you can begin to use it to build your TRAX. Note that you can also use the ACE to establish hyperlinks from the relevant citation entries in your TRAX to the source materials that underlie the citations you are activating. These source materials will already have been loaded into the ACE to ease activation and hyperlinking. Should you happen to forget to provide us with a source file (or wish to activate a citation you had not anticipated activating when you sent us your materials), you will also be able to upload individual source files from your personal computer while working within the ACE. Please make sure that they are correctly named (see below).

As we noted, we understand that not everyone is comfortable learning new software. If using the ACE to create your TRAX would be a significant obstacle to your making progress with your active citation pilot project, we are happy to make alternative arrangements. Please contact QDR staff if you would like help creating your TRAX.

[26] Please contact QDR staff to access the ACE in advance of QDR's launch.

[27] We are drafting template letters for scholars developing QDR pilot projects to send to presses and journals to request permission to post an electronic version of their publication to QDR, and ideally to secure a manipulable PDF of the chapter or article.

A. Additional Information on Providing Underlying Sources

Underlying sources may be in various digital formats: alpha-numeric (e.g., word-processed, databases, spreadsheets), audio (digital and non-digital), graphical, video, and photographic (digital and non-digital). While the ACE can only accommodate PDFs and a single image format (.png) currently, if some of the citations that you activate refer to *non-sensitive* sources in other formats, please submit them to us using Dropbox and we will see that they become hyperlinked to your TRAX through alternative means.

If you only possess a hard copy version of a source (e.g., newspaper clippings, typed or hand-written notes, photographs, documents), please PDF it (and upload it to Dropbox) so that you can use the ACE to establish a hyperlink to it from your TRAX. Please note that any text that you scan will appear online as you scanned it, so please do your best to scan all text right-side-up, and ensure that it is readable. If the source's pages are numbered, each page you scan includes the number.

For sources that are pages from a book, please be sure to include the title page and the copyright page in addition to the pages that you have cited. For sources that are pages from an article, please include the first page, or the page on which the journal, volume, and issue are indicated.

Please use the following convention to name the PDF'd versions of your sources:
FirstElementForSourceInBibliography_YearOfSource_YourLastName

To take an obvious example, for a single-authored secondary source, the file name would be:
AuthorOfSource_YearOfSource_YourLastName

So if Professor Hammond sending us pages 26-29 of an article by Theda Skocpol published in 1985, the file name would read:
Skocpol_1985_Hammond

As is typical in citing academic work, if two or more works by the same author published in the same year are referenced, please add "a" after the year for the first, "b" after the year for the second, etc. Again, file names should track the way in which sources are referenced in the bibliography of the publication being activated.

VI. Contacting

The pilot projects fall into three thematic groups, each with a different primary liaison, listed below. Scholars interested in finding out more about active citation can send inquiries to any one, or all three scholars.

- Andrew Moravcsik (amoravcs@Princeton.EDU) is primary liaison on a group of pilots focusing on Comparative Politics, American Political Development, and International Relations.

- Colin Elman (celman@maxwell.syr.edu) is primary liaison for scholars developing pilots who attended the World War I conference at Syracuse University in April 2012.
- Diana Kapiszewski (dk784@georgetown.edu) is primary liaison for pilots in Comparative Politics, with a particular focus on Latin America and comparative Law and Courts.

Bibliography

American Political Science Association. 2008. *Guide to Professional Ethics in Political Science*. (Changes proposed by the Committee on Professional Ethics, Rights, and Freedoms, July 2012 available at: <https://www.apsanet.org/content_2483.cfm>.)

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

Lustick, Ian. 1996. "History, Historiography, and Political Science: Multiple Historical Records and the Problem of Selection Bias," *American Political Science Review* 90(3): 605-618.

Mahoney, James. 2012. "The Logic of Process Tracing Tests in the Social Sciences." *Sociological Methods & Research* 41(4): 570-597

Moravcsik, Andrew. 2010. "Active citation: A precondition for replicable qualitative research." *PS: Political Science and Politics* no. 43 (1):29-35.

Moravcsik, Andrew. 2012a. "Active Citation and Qualitative Political Science." *Qualitative and Multi-method Research (Newsletter of the Qualitative and Multi-Method Research section of the American Political Science Association).* Spring: 33-37.

Moravcsik, Andrew. 2012b at http://www.princeton.edu/~amoravcs/library/active_citations.docx.

Qualitative Data Repository. ND. "Guide to Sharing Qualitative Data."

Waldner, David. 2012. "Process Tracing and Causal Mechanisms." In Harold Kincaid, Ed., *The Oxford Handbook of Philosophy of Social Science*. New York: Oxford University Press, pp. 65-84.