Transparency for Text-Based Sources: From Principles to Practice
Perspectives on Politics
Supplementary Materials: Appendix A-F

February 2022

Nikhar Gaikwad, Columbia University
Veronica Herrera, UCLA

Appendix A: Coding Procedures for Review of Existing Transparency Practices

We reviewed every article published in the *American Political Science Review*, the *American Journal of Political Science*, *World Politics*, *Perspectives on Politics and Security Studies*[1] published every other year over a six-year period (2008, 2010, 2012, 2014, 2016, 2018), which included a total of 1,120 articles. Of these, 160 were empirical articles using qualitative methods relying on text-based sources. We excluded articles that used solely quantitative methods or that were mixed quantitative-qualitative methods with qualitative methods playing a very minor role. We were inclusive in coding the use of qualitative methodology; for example, articles that were quantitative but also used illustrative vignettes or background cases were included. We excluded articles whose primary or sole empirical basis were interviews, participant observation or ethnography, and we also excluded articles that were not empirical, (e.g., political philosophy). We then selected empirical articles that used text-based sources as the foundation for their empirical claims. We included articles that used at least five text-based sources as units of analysis (at the unit level, e.g., a constitutional article, as opposed to the source itself, e.g., a constitution) as their empirical foundation. That is, five text-based sources "TBS" could be five passages from a constitution, even though the constitution is one source; the unit, here, is the passage, not the constitution.

We found that for the five years reviewed, 160 empirical articles used qualitative methods as well as text-based sources as a substantial component of the methodological approach and evidentiary basis of claims. The fact that less than 15% of articles surveyed made it into our sample likely reflects the publication choices of top political science journals; a similar percentage would have likely resulted had we chosen a different group of top journals. Of these 160 articles, only 28 included supplemental material or appendices of some sort.[2]

---

[1] *Regarding the inclusion of Security Studies:* We initially reviewed APSR, AJPS, WP, and PoP; subsequently, we selected SS precisely because it has historically been receptive to qualitative work and allows for submissions of up to 15,000 words, allowing comparisons with the transparency practices of the other four journals. Our selection method also does not capture transparency practices in other types of publications, such as book chapters or books, the latter of which is likely to provide more opportunities for research transparency due to higher word count limits.

Our analysis shows that of the 20,894 total citations of scholarly sources found in the 160 articles we surveyed, only 43 percent provided page numbers for in-text citations or notes. The average masks significant heterogeneity across journals, however. In SS, which we selected anticipating stronger transparency norms in qualitative research, 67 percent of citations contained page numbers for in-text citations or notes; excluding SS, in the other four journals, only 22 percent of citations contained page numbers. In addition, 39 percent of those citations with page numbers were in cases of directly quoted text; this proportion rises to 51 percent when SS is excluded from the analysis.

[2] We note that other forces may have contributed to trends in transparency documented here, such as copyeditors asking authors to delete page numbers where direct quotations are not referenced, lack of clarity for norms regarding

One person coded all of the articles after corroborating intercoder reliability for a small subset of articles with a second researcher; we did not conduct further intercoder reliability tests. We used SCOPUS as our search engine. Below we outline the article selection and coding procedures we used through two key steps. In the spirit of transparency, we also include notes from our coder as they proceeded through this research exercise.

Step 1: Article Selection Process

- First generate search for all articles for the year. Scopus: Journal Title (Source Title), year (2018, 2016, 2014, 2012, 2010, 2008). Limit search to articles only. Each article in the relevant year needs to be examined through abstract, and then opened. Determine whether it belongs in "codeable" pile. If it does, save a copy, and enter details in spreadsheet (author year). Download and also save appendices and supplementary materials (same author year, plus appendices and supplementary materials).
- To be entered into codeable count, article needs to:
    a. Use text-based sources in a significant way as evidentiary base through a qualitative analysis.
        i. Thus, primarily quantitative studies can be excluded.
        ii. Illustrative vignettes or background cases that are important parts of evidence count if they rely on text-based sources (as opposed to e.g. interviews)
    b. Be empirical (as opposed to pure theory). (Hints: Does article have a methods section? Does article propose to analyze texts or sources in systematic way for descriptive or causal argument? Is article written by a political theorist or claims to develop political theory? Articles by political theorists that include empirical material using text-based sources as a primary form of evidence were included).
    c. Use (some, specifically > 5) text-based sources as evidence for argument being made. Note that the count is at the unit level rather than the actual source itself. (e.g., 5 TBS "sources" could be five passages from a constitution, even though the constitution is one source. The unit, then is the passage, not the constitution in terms of the count used to determine codeable status). When in doubt, count the source. Examples of TBS types: Secondary sources[3] (when being used as

---

referencing source locations in online journals, style manuals not updating citation standards for social media sources, and other publishing pressures.

[3] *A Note on Secondary Sources:* We define secondary sources as scholarly journal articles, books and book chapters that are typically written by scholars for an academic audience. We follow the discipline of history and define secondary sources as those created by someone who did not experience first-hand or participate in the events or conditions being described or analyzed. While secondary sources can be viewed as text-based sources broadly defined, we excluded these sources from our coding analysis for source location, production, selection, analysis and access except for the coding we performed for determining page number citation percentages, where we only included secondary sources or scholarly sources such as journal articles, books and book chapters (i.e., Table 3).

However, sometimes secondary sources can be used in ways that approximate the use of primary sources in qualitative empirical research in political science. For the purposes of our coding of source types presented in Table 1, we included two types of secondary sources that were used for analysis in qualitative empirical research. These fell under three broad categories, although the first was the most common 1) books and journal articles that were *historical interpretations* (e.g. historical treatises, military histories, histories of presidential administrations, etc.), 2) *policy analysis* from a scholarly source that were descriptive pieces about a particular policy (e.g. Bush doctrine for empirical evidence in main body of paper to support argument being made, not in literature review); Archival material; Government data; Newspaper articles; Organization's publications; NGO reports; Resolutions, Accords, Constitutions; Multimedia (e.g., video, audio clips, maps).

Step 2: Code Article along five categories:
Location, Production, Selection, Analysis, Access

- In the spreadsheet, make a note of the primary format for transparency communication: in-text, footnote/endnote, appendix, or methods statement.
- Highlight in blue the places where transparency communication is happening within the article PDF wherever relevant or feasible. Save a copy of the highlighted version.

1) Source Location.
    a. Did the primary TBS being used get discussed in terms of where it is located such that another person could find it? Code 0 or 1. Especially regarding foreign language items, was it in original language so the source could be located by others? When it is translated and the link is broken it is very difficult to find the source). This may vary across data source type, for example. Source location was searched within-text, in footnotes/endnotes, appendices, and methods statements. Additional points to consider:
        i. *Newspaper articles*: Full title, full date, html link. (Author may not be necessary if source can be found without author). Needs html or other indication to be complete (especially if you search online and can't find the source—oftentimes these links get broken).
        ii. *Archival materials*: Where in the archive, box #s, etc., and location of the archive?
        iii. (*Organizational) Resolutions, Mandates,* etc. need names, titles, location within document (e.g. pg # or para, or article), but also information about where they are physically located
        iv. *Online data sources.* Requires names, html links, dates accessed.
    b. Notes from our coder on how they managed the coding process for this category: "I would first look at the type of TBS being used. Secondary sources were straightforward to determine location. For other documents, I would look in footnotes or bibliography. URLs were the easiest way of determining location but I would also look for physical location and then see if the physical location of the text was specific enough for another researcher to find as well. So, for example, just mentioning a city, would not qualify. Physical texts would need an easily mappable location with perhaps a building, stack number, or reference number within a specific set of archives so other researchers could easily access the source without having to dig to find it."

2) Source Production
    a. Did the primary TBS being used get discussed in terms of how it was produced, by whom, etc.? Beyond just citation, did the authors give background information

---

foreign policy) and 3) *Foreign language scholarly analysis* on a wide range of specialized issues that political scientists drew on to help build case studies or other qualitative analysis.

as to what person or institution actually made the document? Did they discuss anything about what circumstances surrounded the making of the document, what biases or agendas were being furthered by the making of the document, or the level of trustworthiness, expertise, etc., related to the creators of the document? If archival material, was any information on the archive and how it is curated included? Code 0 or 1.

b. Notes from our coder on how they managed the coding process for this category: "Production was rarely discussed by any of the authors. I gave production a 1 if they gave useful background information on the organizations that produced the documents or if they gave information on the context of the document. An example of a 1 is Valdez 2016 on p. 19. The author gives the background context for the passage of the law before writing about the law itself. Through this contextualization, the author gives the audience a lens through which to understand the law before using it as evidence. Another example of this is Gotham 2012 who gives context on why FEMA and DHS were created before using documents produced by the two agencies. I found that if production is not mentioned before the sources are used, then it likely won't be mentioned. I checked footnotes and supplementary material to confirm this."

Valdez, Inés. 2016. "Nondomination or Practices of Freedom? French Muslim Women, Foucault, and The Full Veil Ban." *American Political Science Review* 110(1): 18–30.

Gotham, Kevin Fox. 2012. "Disaster, Inc.: Privatization and Post-Katrina Rebuilding in New Orleans." *Perspectives on Politics* 10(3): 633–646.

3) Source Selection

a. Did the primary TBS being used get discussed in terms of how it was selected by researcher for use in this study? Does author explain what made them use this source instead of another? Does the author discuss anything about why they are using the source? Code 0 or 1.

b. Notes from our coder on how they managed the coding process for this category: "Froio 2018 p. 701 in *Perspectives* is a good example of selection. The author gave a clear roadmap of what made them include certain sources and exclude others. After checking the supplementary materials and footnotes, I would scan the article and see if the author provided any background information on their sources. If they did, I would decide if it either fell under production or selection. For selection, the author would have to go beyond just contextualizing the source and actually describe how they arrived at using the source to make their claim. Selection was given a 1 if you could easily tell why the author chose to use the source they did over other relevant sources."

Froio, Caterina. 2018. "Race, Religion, or Culture? Framing Islam between Racism and Neo-Racism in the Online Network of the French Far Right." *Perspectives on Politics* 16(3): 696–709.

4) Source Analysis

- Note: This category was the most difficult to code, it was a more subjective category than the other four.
    a. Did the primary TBS being used get discussed in terms of how the source supports the claims being made? Did this occur consistently throughout? More than just a simple citation, does the author draw the line from the source to their claim? Do they do this consistently with most of the main text-based sources being coded in our spreadsheet? Code 0 or 1.
    b. Notes from our coder on how they managed the coding process for this category: "I looked to see if the author first explains what the original TBS was arguing, or what it was saying, before the author used it to support their argument. Did the author describe what the original source was saying and then link it to the point they were making? It was easy to eliminate an article if the author seemingly used citations but then gave no evidence or explanation to what part of the author's argument the citation is supporting. If direct quotes were used, I looked to see if they explained the quotes and directly applied them to their argument (Green 2008, *Perspectives*; Baccini & Koenig-Archibugi 2014, *World Politics*). Meyer 2016 is also an example of good analysis; the author makes a point, uses an in-text citation, and then expands upon the point or pulls in a direct quote that explained the use of the citation in the footnotes. Analysis was given a 1 if there was little room for interpretation on how the claim being made is supported by the TBS, direct quotes were the most useful in determining this category. Another key factor for receiving a 1 was consistency; if the author consistently explained with direct quotes or showed how the source used supported the claim, then they would receive a 1."
    c. As noted above, coding for analytical transparency was the most difficult of the five categories because of the many different epistemological communities and research traditions represented across the articles surveyed, and the diverse manner in which text-based sources were employed in articles. Yet we believe that it is a possible to create a decision rule that captures this category so that it is codable in a large review such as the one we undertook in this paper. The decision rule that we chose was that an article had to have had at least three (3) instances of analytic transparency connected to a text-based source in order to be coded as demonstrating analytic transparency. However, in order to create sensitivity bounds for this analysis, we randomly selected and reviewed 1/3rd of our sample in order to ascertain results from a more permissive decision rule.[4] In particular, under the more permissive decision rule, an article was coded as demonstrating analytic transparency if it had just one (1) or more instances of analytic transparency connected to a text-based source. When an article was borderline in terms of being coded for analytic transparency, we gave it a "1". We then created sensitivity bounds by studying the number of articles that were coded as demonstrating analytic transparency and by extrapolating this proportion to the entire sample of 160 articles. Under this more permissive coding, the total number

[4] We randomized the selection by starting with the first article in our sample and then selecting every fifth article across each journal in the order *APSR, AJPS, Perspectives, World Politics, Security Studies*.

of articles receiving a "1" for analytic transparency for the 160 articles would increase to 33, or 20% of the sample. That is, while our original coding resulted in 12% of articles in the entire sample being classified as demonstrating analytic transparency (see Table 2), the more permissive coding scheme resulted in 20% of articles in the entire sample being classified as demonstrating analytic transparency, which is still quite a low proportion. This permissive coding scheme is an upper bound since articles merely needed to have only one (1) instance of analytic transparency connected to a text-based source to get classified as demonstrating analytic transparency.

Baccini, Leonardo and Mathias Koenig-Archibugi. 2014. "Why do States Commit to International Labor Standards? Interdependent Ratification of Core ILO Conventions, 1948–2009." *World Politics* 66(3): 446–490.

Green, Elliott. 2008. "Understanding the Limits to Ethnic Change: Lessons from Uganda's 'Lost Counties'." *Perspectives on Politics* 6(3): 473–485.

Meyer, Brett. 2016. "Learning to Love the Government: Trade Unions and late Adoption of the Minimum Wage." *World Politics* 68(3): 538–575.

5) Source Access
    a. Did the primary TBS being used get shared or have excerpts from the source shared? For quotes from the same source type, were authors consistently using quotes that encompassed most of the point being made? Must be recurring. If URL is provided, does URL bring you directly to source (or source excerpt)? Should not have to dig at all, source should be there, in its entirety, through link provided, or in the appendix or on the author's website. Code 0 or 1.
    b. Notes from our coder on how they managed the coding process for this category: "To determine if something deserved a 1 or a 0, I would first look at the section of the paper that used qualitative evidence (see Column E in the workbook). I would take note of the citations used and then go to either the footnotes, end notes, supplementary material, or bibliography to see if there was a link to the full source or if the source was replicated. Even if there was a URL, I would then confirm that it led to the source and was still an active website and not a broken link."

Appendix B: Full Citations for the Qualitative Transparency Deliberations Board Posts

For a detailed description of the QTD background, objectives, and process, see: www.qualtd.net/page/about/ and also Jacobs et al (2021, 172–76). A complete archive of the online deliberations is available on Harvard Dataverse: https://doi.org/10.7910/DVN/SWVFV8. See Stage 2-Working Group II.1. Text based sources.

Full Citations

Kristen Harkness, "What might qualitative data access look like?," QTD Discussion Board, April 27, 2016.

Shamira Gelbman, "Documenting use of text-based or non-text-based sources," Discussion forum II.1, QTD Discussion Board, December 1, 2016.

Jane Mansbridge, "Benefits and Costs of Increasing Transparency for Text and Non Text Based Sources," Discussion forum II.1, QTD Discussion Board, October 26, 2016.

Jacques Hymans, "Benefits and Costs of Increasing Transparency for Text and Non Text Based Sources," Discussion forum II.1, QTD Discussion Board, December 13, 2016.

Chloe Thurston, "Active citation versus the meaty footnote," QTD Discussion Board, May 17, 2016.

Guest, "Documenting use of text-based or non-text-based sources," Discussion forum II.1, QTD Discussion Board, December 9, 2016.

Margaret Keck, "No place for my work in this debate," QTD Discussion Board, April 8, 2016.

Giovanni Capoccia, "Data access and 'right to first use' of newly collected quantitative data," QTD Discussion Board, May 19, 2016.

Sheena Greitens, "DA-RT: effect on graduate training" QTD Discussion Board, April 20, 2016.

Sam Handlin, "Benefits and Costs of Increasing Transparency for Text and Non Text Based Sources," Discussion forum II.1, QTD Discussion Board, October 17, 2016.

Amy Poteete, "Benefits and Costs of Increasing Transparency for Text and Non Text Based Sources," Discussion forum II.1," QTD Discussion Board, January 1, 2017.

Appendix C: Additional Results from Review of Published Articles

Table 1: Number of Published Articles using Text Based Sources (2008-2018)

	American Political Science Review	American Journal of Political Science	World Politics	Perspective on Politics	Security Studies
2018	2	1	8	9	14
2016	7	0	10	5	11
2014	1	1	4	2	13
2012	2	1	12	4	11
2010	0	0	6	4	11
2008	1	0	5	3	12
Total	13	3	45	27	72

Note: We only included articles whereby the methodology (form of analysis) used was primarily qualitative. Within this subset of articles, we excluded articles whereby the predominant evidence base was not text based sources (e.g. interviews, ethnography). For qualitative research using text-based sources, we excluded those that only used secondary sources.

Appendix D: Additional Illustrative Examples of Transparency in Practice

In the main manuscript, we provide a detailed analysis of three published articles that implement transparency practices across our five dimensions of transparency. In this section, we provide several additional examples from scholarship drawing from different subfields and corners of the discipline in both journal articles and scholarly books in order to further illustrate that many scholars are already practicing forms of transparency about text-based sources and how they do so.

First, transparency about source location helps other researchers locate data and evaluate it, expanding the scope and reach of one's research. For appropriate assessment of research based on texts, it is necessary to provide detailed information about where sources are located if they are publicly available. Providing detailed page numbers when sources are used as empirical evidence, or providing specific archival location information down to the last identifier, is an important first step toward "findability" and external assessment. For example, Kimberly

Johnson provides excellent identifier information in her archival research on African American poll-tax registration and voting when she specifies, "This article draws from the following archival sources: Luther Porter Jackson Papers, 1772-1960, Accession No. 1952-1, Special Collections and Archives, Johnston Memorial Library, Virginia State University, Petersburg, VA, hereafter referred to LPJ papers." She also references her sources down to the last possible identifier available—for example, "'New Canton Voters Association Meeting (Minutes),' May 1946, Folder 548, Box 19, LPJ" (2017, 220, fn 13, 231 fn 73).

Similarly, Sarah Goodman goes beyond citing a newspaper title and date, a practice which has become all too common. In her work on citizenship and immigration, Goodman cites her source as "Philip Johnson, 'Migrants Face New "Britishness" Test.' *Telegraph*. December 5, 2006" (Goodman 2012, 687, fn 88). By providing the newspaper article's author, the title in the original language, the newspaper name, and the exact date, it is more likely that others can find the source in the future even if original URLs are broken. In making citations, researchers should go beyond what reviewers or editors require for specifying location information for text-based sources. They should consider what subsequent readers will need to know about where publicly available sources are located in order to evaluate or build on their scholarship.

Second, specificity about a source's production allows readers to better understand the quality of the data being used and its appropriateness for a particular evidentiary claim. For example, in *Networks in Contention: The Divisive Politics of Climate Change*, Jennifer Hadden explains how she was "subscribed to the private internal listservs of Climate Action Network Europe and Climate Action Network International, as well as to the public lists of Climate Justice Alliance and Climate Justice Network. During the time period of this study, I collected more than 10,000 emails through these channels, which kept me exceptionally well informed about the political process and the workings of these coalitions. As all these emails are confidential, I do not discuss their contents except when they exist in the public sphere or have been discussed in on-the-record interviews" (Hadden 2015, 194). By disclosing the private origins of her source, Hadden adds credence to her evidentiary claims regarding the inner workings of a climate activism organization. While this work may not be replicable due to the private nature of the source, the information given on the source's production nevertheless boosts the author's claims.

Specificity about source production can occur even for informal archives where authors may have permission to access archives but not share their contents. In his study on slum leadership and public service provision in India, Adam Auerbach photographed and digitally organized materials curated by local leaders in informal archives and described how these materials were produced and maintained. Auerbach explained how local leaders chose to preserve some materials while others were destroyed or deteriorated over time, stressing that "the decision to maintain materials is a choice that, like documentation, is non-random" (2018, 360). Auerbach described how photographs, petitions, letters from officials, party materials, and community meeting notes provided key insights, but also how they were often incomplete. Individuals were less likely to share documentary information that might cast them in a negative light, and local leaders' were more apt to keep particularly well-maintained written records of citizen petitions that documented how they as leaders engaged in problem-solving efforts (Auerbach 2018, 351, 354). Furthermore, important events were left out of the informal archival materials when they related to the Hindu minority who were socially isolated in the Muslim community (Auerbach 2018, 359–60). Auerbach's description of how his text-based sources were produced help readers understand their evidentiary value as well as their limitations. It also demonstrates the importance of triangulating with other sources, such as interviews and ethnography.

Third, specificity about a scholar's source selection process provides insights into her methodology, the degree to which a source was "cherry-picked," and the context for the different types of biases that may present. For instance, Alisha Holland uses newspaper articles to measure attitudes about public perceptions on street vending in Lima and Bogotá in her book *Forbearance as Redistribution: The Politics of Informal Welfare in Latin America* (Holland 2017, 164–68, 334–37). In Chapter 4 and in "Appendix B: Coding Rules for Campaign Platform, Newspaper and Administrative Sources," Holland describes how she constructed a database of these articles. She specifies the date range of the search and offers several categories of news articles selected (news items, letters to editors, short notes, platforms, speeches, interviews). She then coded content based on whether or not it was sympathetic to the practice of street vending—with discussions regarding unemployment, displacement, criminalization of poverty coded as sympathetic, and discussions regarding public space, business competition, public safety risk, and crime coded as unsympathetic. Furthermore, Holland explains why she chose to code newspaper articles and not rely on public opinion polling (it didn't exist), when she used online or physical newspapers, why she chose the newspapers she chose, and her triangulation strategies with information from campaign platforms and administrative sources.

In a similar vein, Ruth Rubin identifies what archival sources were available to her and how representative they are in her book, *Building the Bloc: Intraparty Organization in the US Congress* (2017). Rubin selected organizations and their archives for study in order to cover a range of contextual variation. She argues that these archives, not the official Congressional Record, are the most appropriate source on intra-party organization because the researcher can "follow members of Congress off the floor and behind closed doors" (Rubin 2017, 26). The book also notes source limitations, stating that, "These data are far from perfect. Members of Congress do not preserve all of their correspondence and may not be entirely forthcoming, even in confidential letters" (Rubin 2017, 26). Rubin notes the absence of data in many places, including how the "lack of formal records" among Republican Senate insurgent organizations "complicates our analysis of the lawmakers' organizational activity. In the absence of organizational records, we must rely first and foremost on individual records" (Rubin 2017, 72). However, she leverages these data limitations to support her argument, proposing that, "We can view the organization's lack of formal record keeping as a data point in and of itself" (Rubin 2017, 72). Rubin also notes how a broader evidentiary base lends credence to the cited source. For instance, she identifies the "strong archival evidence" for press coordination between House and Senate Republican insurgents, then cites a particular letter emblematic of this evidence base (Rubin 2017, 47, fn 62). The book's appendices include a summary of all archival collections she consulted, lists of all the meetings whose minutes she reviewed, and excerpts of key documents.

Fourth, specificity about how a source is interpreted and analyzed helps readers assess why the sources being provided are indeed evidence for the author's claims. For example, in his article on how converts to Protestant Christian sects influenced democracy-building in Asia, Robert Woodberry makes the claim that "Christian converts published the first privately printed Japanese-and Korean-language newspapers" (Woodberry 2012, 250, n. 17). He provides detailed reasoning in a "meaty footnote" that helps substantiate his claim by noting, alongside additional citations, that,

> The first privately printed Japanese-language newspaper was printed by Hamada Hikizō /Joseph Heco, a Protestant who had worked with missionary printers, and Kishida Ginkō, a student of the missionary Joseph Hepburn….An earlier government-printed paper was a translation of a Chinese-language missionary newspaper, minus the religious content. The

> Japanese government distributed it to a small number of high government officials as a way to monitor the outside world. It was not available to the public. The first privately printed Korean-language newspaper (the Independent) was edited by Philip Jaisohn/So Chaep'I, a Protestant teacher at a mission school. Missionaries encouraged him to publish it, provided the trained printing staff free of charge, and continued printing the paper after he fled Korea.…

In providing this explanation, Woodberry strengthens the validity of his claim and signals scholarly rigor. Word count limits at various publications have made analytical transparency a more challenging task for qualitative researchers. Nevertheless, describing the analytical route through which authors link their claims to sources is critically important for augmenting transparency.

An added benefit to this transparency practice is that it establishes a trail for researchers to follow and "retrace" their steps throughout the research process, especially when data archiving is added to the mix. Producing a transparency appendix during the actual research process is also less burdensome than scrambling to create one in the last stages of publication (Saunders 2014, 696). For example, Veronica Herrera notes that in her experience, developing a transparency appendix "in the moment," (rather than retroactively) facilitates the research and writing process by providing a built-in organizational structure.[5]

Finally, sharing excerpts or full replications of text-based sources provides a plethora of benefits. Excerpts from sources can help other researchers understand why a scholar is using certain language to make key claims and how that language is linked to evidence from original sources. For example, in his book *The Rights Revolution: Lawyers, Activists, and Supreme Courts in Comparative Perspective*, Charles Epp argues that Canada underwent a major rights

[5] Herrera also noted that online data archiving can be a highly useful system for keeping files organized, especially if one's computer and backup crash, which happened to Herrera in 2015. She was able to restore her files because they were stored in a QDR transparency appendix (at the time a pilot project as an Active Citation Compilation, now an Annotation for Transparent Inquiry) (Herrera 2015).

revolution for the rights of the accused in the 1980s, and that NGOs and other organizations helped provide support for this rights revolution. He shows how leading Canadian civil liberties associations participated in key revisions to the Canadian Constitution's Charter of Rights and Freedoms' wording on search and seizure procedures, exclusion of evidence, detention and imprisonment, and the right to counsel and jury trials (Epp 1998, 188). Epp's claims are strengthened through his thirty-two line endnote with detailed excerpts of the before and after wording of provisions in the Charter regarding each of these legal rights (Epp 1998, 276–77, n. 59).

Providing a full source for a document that is in the public domain but that is hard for other researchers to find can constitute an academic public good. For example, in their study on social accountability strategies for healthcare and environmental policy in Colombia, Veronica Herrera and Lindsay Mayka (2020) create a text-based source appendix that provides the full texts of publicly available sources cited in their paper, including high court rulings, reports from the offices of the Attorney General and Ombudsman, and eighteen newspaper articles. Even though these materials are available to the public, they were gathered during field research and would not be easy for others to find without similar fieldwork. Although most researchers likely have a personal document management system for their individual workflow, organizing documents in a way that prioritizes transparency and access for other researchers will have additional benefits. Authors will be less likely to cut corners and more likely to add additional contextual information to their source notes when they anticipate how others will access this information to assess the project's contents.

Ultimately, "transparency cannot substitute for good research" (Saunders 2014, 694). However, research transparency does provide important benefits to the field—such as signaling scholarly rigor, improved communication of evidence and findings, the development of research subfields and topics as more primary sources and data become widely available, improved qualitative research training for graduate students, and increased ease of scholarly exchange.

Appendix E: Copyright and Fair Use Resources

APSA Style Manual's Summary of Fair Use Doctrine
"The fair use doctrine, which has developed over time, is seen as one of the corner stones of free expression in the United States. Fair use limits copy right to balance the interests of copyright holders with public interest in the wider distribution and use of creative works… Fair use is determined by (1) the purpose and character of use, whether for commercial or educational purposes, (2) the nature of the copyrighted work, (3) the amount and substantialness of the portion in relation to the whole, and (4) the effect of the use on the potential market for the copyrighted work… The use of an entire literary work in its entirety is hardly ever accept-able. Use that is not fair will not be excused by paraphrasing, as it is considered disguise copying by copyright doctrine (4.89)." (APSA 2018, 7).

Stanford Libraries, "What is Fair Use?"
"In its most general sense, a fair use is any copying of copyrighted material done for a limited and "transformative" purpose, such as to comment upon, criticize, or parody a copyrighted work. Such uses can be done without permission from the copyright owner. In other words, fair use is a defense against a claim of copyright infringement. If your use qualifies as a fair use, then it would not be considered an infringement," Stim (2019, unknown page numbers). For more information, see https://fairuse.stanford.edu/overview/fair-use/what-is-fair-use/.

Websites

Fairuse.stanford.edu (Stanford Libraries, Copyright and Fair Use, Copyright FAQs; Fair Use)

Copyright.gov

Richard Stim's website: dearrichblogspot.com

Appendix Bibliography

APSA. 2018. *Style Manual for Political Science: Revised 2018 Edition*. Washington D.C.: American Political Science Association.

Auerbach, Adam. 2018. “Informal Archives: Historical Narratives and the Preservation of Paper in India’s Urban Slums.” *Studies in Comparative International Development* 53 (3): 343–64.

Epp, Charles. 1998. *The Rights Revolution: Lawyers, Activists, and Supreme Courts in Comparative Perspective*. Chicago: Chicago University Press.

Goodman, Sara Wallace. 2012. “Fortifying Citizenship: Policy Strategies for Civic Integration in Western Europe.” *World Politics* 64 (4): 659–98.

Hadden, Jennifer. 2015. *Networks in Contention: The Divisive Politics of Climate Change*. New York, NY: Cambridge University Press.

Herrera, Veronica. 2015. “Data for: "Does Commercialization Undermine the Benefits of Decentralization for Local Services Provision? Evidence from Mexico’s Urban Water and Sanitation Sector.” Qualitative Data Repository. 2015. http://doi.org/10.5064/F6F769GQ.

Herrera, Veronica, and Lindsay Mayka. 2020. “How Do Legal Strategies Advance Social Accountability? Evaluating Mechanisms in Colombia.” *Journal of Development Studies* 56 (8): 1437–54.

Holland, Alisha C. 2017. *Forbearance as Redistribution: The Politics of Informal Welfare in Latin America*. New York: Cambridge University Press.

Jacobs, Alan M., Tim Buthe, Ana M. Arjona, Leonardo R. Arriola, Eva Bellin, Andrew Bennett, Lisa Bjorkman, et al. 2021. “The Qualitative Transparency Deliberations: Insights and Implications.” *Perspectives on Politics* 19 (1): 171–208.

Johnson, Kimberley S. 2017. “From Politics to Protest: African American Voting in Virginia in the Pre–Civil Rights Movement Era, 1940–1954.” *Studies in American Political Development* 31 (2): 218–37.

Rubin, Ruth Bloch. 2017. *Building the Bloc: Intraparty Organization in the US Congress*. New York: Cambridge University Press.

Saunders, Elizabeth N. 2014. “Transparency without Tears: A Pragmatic Approach to Transparent Security Studies Research.” *Security Studies* 23 (4): 689–98. https://doi.org/10.1080/09636412.2014.970405.

Stim, Richard. 2019. *Getting Permission: Using & Licensing Copyright-Protected Materials Online & Off*. 7th ed. NOLO.

Woodberry, Robert D. 2012. “The Missionary Roots of Liberal Democracy.” *American Political Science Review* 106 (2): 244–74.

Appendix F: Reading Lists on Transparency Issues for Qualitative Researchers

Below we provide sources for further reading on qualitative research transparency. The list is incomplete but nevertheless a good starting point.

Further Reading from Qualitative Data Repository Affiliated Authors (found on QDR website, downloaded on 1.28.22)

Jacobs, Alan M., Diana Kapiszewski, and Sebastian Karcher. 2022. "Using Annotation for Transparent Inquiry (ATI) to Teach Qualitative Research Methods." PS: Political Science & Politics 55 (1): 216–20. https://doi.org/10.1017/S1049096521001335.

VandeVusse, Alicia, Jennifer Mueller, and Sebastian Karcher. 2021. "Qualitative Data Sharing: Participant Understanding, Motivation, and Consent." Qualitative Health Research. doi: 10.1177/10497323211054058.

Kapiszewski, Diana, and Elisabeth Jean Wood. 2021. "Ethics, Epistemology, and Openness in Research with Human Participants." Perspectives on Politics 1-17. https://doi.org/10.1017/S1537592720004703.

Demgenski, Robert, Sebastian Karcher, Dessi Kirilova, and Nic Weber. 2021. "Introducing the Qualitative Data Repository's Curation Handbook." Journal of eScience Librarianship 10 (3): 1207. https://doi.org/10.7191/jeslib.2021.1207.

Kapiszewski, Diana, and Sebastian Karcher. 2021a. "Transparency in Practice in Qualitative Research." PS: Political Science & Politics 54 (2): 285–91. https://doi.org/10.1017/S1049096520000955.

———. 2021b. "Empowering Transparency: Annotation for Transparent Inquiry (ATI)." PS: Political Science & Politics 54 (3): 473–78. https://doi.org/10.1017/S1049096521000287.

———. 2021c. "Introduction: Case Studies in Transparent Qualitative Research." Qualitative & Multi-Method Research 19 (1): 6–9. https://doi.org/10.5281/zenodo.5495548.

Karcher, Sebastian, Dessislava Kirilova, Christiane Pagé, and Nic Weber. 2021. "How Data Curation Enables Epistemically Responsible Reuse of Qualitative Data." The Qualitative Report 26 (6): 1996–2010. https://doi.org/10.46743/2160-3715/2021.5012.

Weber, Nicholas, and Sebastian Karcher. 2020. "Seeking Justification: How Expert Reviewers Validate Empirical Claims with Data Annotations." In Proceedings of the ACM/IEEE Joint Conference on Digital Libraries in 2020, 227–34. JCDL '20. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3383583.3398537.

Weber, Nicholas, Sebastian Karcher, and James Myers. 2020. "Open Source Tools for Scaling Data Curation at QDR." The Code4Lib Journal, no. 49 (August). https://journal.code4lib.org/articles/15436.

Alexander, Steven M., Kristal Jones, Nathan J. Bennett, Amber Budden, Michael Cox, Mercè Crosas, Edward T. Game,... Sebastian Karcher, ..., Nic Weber et al. 2019. "Qualitative Data Sharing and Synthesis for Sustainability Science." Nature Sustainability, November, 1–8. https://doi.org/10.1038/s41893-019-0434-8.

Karcher, Sebastian, and Sophia Lafferty-Hess. 2019. "An Epic Journey in Sharing: The Story of a Young Researcher's Journey to Share Her Data and the Information Professionals Who Tried to Help." IASSIST Quarterly 43 (1): 1–10. https://doi.org/10.29173/iq942.

Karcher, Sebastian, and Nicholas Weber. 2019. "Annotation for Transparent Inquiry: Transparent Data and Analysis for Qualitative Research." IASSIST Quarterly 43 (2): 1–9. https://doi.org/10.29173/iq959.

Fenner, Martin, Mercè Crosas, Jeffrey S. Grethe, David Kennedy, Henning Hermjakob, Phillippe Rocca-Serra, Gustavo Durand, ..., Sebastian Karcher, ... et al. 2019. “A Data Citation Roadmap for Scholarly Data Repositories.” Scientific Data 6 (1): 28. https://doi.org/10.1038/s41597-019-0031-8.

Mannheimer, Sara, Amy Pienta, Dessislava Kirilova, Colin Elman, and Amber Wutich. 2018. “Qualitative Data Sharing: Data Repositories and Academic Libraries as Key Partners in Addressing Challenges.” American Behavioral Scientist, June, 0002764218784991. https://doi.org/10.1177/0002764218784991.

Elman, Colin, Diana Kapiszewski, and Arthur Lupia. 2018. “Transparent Social Inquiry: Implications for Political Science.” Annual Review of Political Science 21 (1): 29–47. https://doi.org/10.1146/annurev-polisci-091515-025429.

Kirilova, Dessi, and Sebastian Karcher. 2017. “Rethinking Data Sharing and Human Participant Protection in Social Science Research: Applications from the Qualitative Realm.” Data Science Journal 16 (September). https://doi.org/10.5334/dsj-2017-043.

Karcher, Sebastian, Dessislava Kirilova, and Nicholas Weber. 2016. “Beyond the Matrix: Repository Services for Qualitative Data.” IFLA Journal 42 (4): 292–302. https://doi.org/10.1177/0340035216672870.

Elman, Colin, and Colleen Dougherty Burton. 2016. “Research Cycles: Adding More Substance to the Spin.” Perspectives on Politics 14 (4): 1067–70. https://doi.org/10.1017/S1537592716002991.

Elman, Colin, Diana Kapiszewski, and Dessislava Kirilova. 2015. “Learning through Research: Using Data to Train Undergraduates in Qualitative Methods.” PS: Political Science & Politics 48 (1): 39–43. https://doi.org/10.1017/S1049096514001577.

Elman, Colin, and Diana Kapiszewski. 2014. “Data Access and Research Transparency in the Qualitative Tradition.” PS: Political Science & Politics 47 (01): 43–47. https://doi.org/10.1017/S1049096513001777.

Kapiszewski, Diana, and Dessislava Kirilova. 2014. “Transparency in Qualitative Security Studies Research: Standards, Benefits, and Challenges.” Security Studies 23 (4): 699–707. https://doi.org/10.1080/09636412.2014.970408.

Lupia, Arthur, and Colin Elman. 2014. “Openness in Political Science: Data Access and Research Transparency.” PS: Political Science & Politics 47 (01): 19–42. https://doi.org/10.1017/S1049096513001716.

Elman, Colin, Diana Kapiszewski, and Lorena Vinuela. 2010. “Qualitative Data Archiving: Rewards and Challenges.” PS: Political Science & Politics 43 (01): 23–27. https://doi.org/10.1017/S104909651099077X