Can We Do Better? Replication and Online Appendices in Political Science

Jonathan Grossman and Ami Pedahzur

Replicability in political science is on the rise, as disciplinary journals have been placing a growing emphasis on data access and research transparency (DA–RT) practices and policies. As a result, nearly every article that is published today in leading political science journals offers an online appendix that includes data, code, and methodological explanations necessary for replication. While these developments are laudable, many appendices still do not enable satisfactory replication because they are inaccessible, compartmentalized, and difficult to understand. In this article and in its accompanying online appendix, we demonstrate this problem and make the case for more accessible and comprehensive appendices whose contribution can fulfill and go beyond mere replicability. We propose several ways in which authors and journals can produce better appendices, namely, by making appendices more intuitive, integrated, and standardized, and by choosing an adequate online platform on which to create and host the appendix.

The study of politics has entered the age of transparency. Following the replication crisis in the social sciences, more and more political science journals instruct authors to make their research fully reproducible upon publication. Ideally, every evidence-based work should meet the replication standard, according to which readers should have all the information needed to replicate a study's results without requiring any additional data (King 1995, 444). While we may never fully achieve this replicability utopia (see Gibson 1995), we are approaching it—today, virtually every evidence-based article in leading political science journals offers a methodological appendix, a downloadable repository of replication data, or both. Over the next few years, an increasing number of journals are expected to adopt these norms.

Despite the growing prevalence of downloadable replication materials, these guides rarely meet their intended goals. Common problems include the absence of crucial elements for replication as well as inaccessible or disorganized files. These challenges often prevent full replication. As Arthur Lupia and Colin Elman (2014, 22) emphasize, openness is essential for an effective dialogue between members of different sub-disciplinary and methodological communities within political science (refer to online appendix Annotation 1, henceforth "OAA #1").[1] Accessibility and user experience design, in turn, are prerequisites for real openness. They can make the difference between *primary replication*, in which scholars are able to fully replicate the entire course of a study, and *secondary replication*, which only entails replicating the statistical analysis based on data and code shared by authors (OAA #2).

In this article, we propose methods for providing accessible and transparent replication data in political

A list of permanent links to Supplemental Materials provided by the authors precedes the References section.

**Data replication sets are available in Harvard Dataverse at: https://doi.org/10.7910/DVN/VXZZ0J*

Jonathan Grossman is a postdoctoral fellow at the Leonard Davis Institute for International Relations, the Hebrew University of Jerusalem (jonathan.grossman@mail.huji.ac.il).

Ami Pedahzur is the Ralph W. Yarborough Centennial Professor of Liberal Arts at the Department of Government, The University of Texas at Austin (pedahzur@austin.utexas.edu).

The research for this paper was supported by the Leonard Davis Institute for International Relations at the Hebrew University of Jerusalem. An earlier version of this paper was presented at the Leonard Davis Institute Research Seminar in January 2020. We thank the participants of the seminar for their constructive comments. We are especially indebted to Yoram Haftel and Raelene Camille Wyse for their excellent suggestions.

doi:10.1017/S1537592720001206

science. We take a holistic approach to such data, viewing all types of them (methodological explanations, contextual information, charts and tables, constructed datasets, raw data files, software syntax, and so forth) as integral parts of a single unified and coherent digital body of knowledge whose different components correspond with one another and with the main article. For the sake of brevity and simplicity, we refer to this ecosystem as the *online appendix*. First, we review the current state of transparency in political science. Second, we demonstrate how the inaccessibility and incomprehensibility of appendices hinder openness and replicability. Third, we discuss key principles and best practices that can improve appendices' impact and accessibility.

The Quest for Transparency in Political Science

The transparency movement in political science reached a critical mass in the early 2010s (Golder and Golder 2016, 2–5; Lupia and Elman 2014, 19–21). Throughout this decade, a series of formal and informal discussions and deliberations concerning transparency took place in political science conferences, committees, journal symposia, and online forums, resulting in various reports and recommendations, as well as relevant criticism (OAA #3). The movement culminated in 2012 with the amendment of the APSA Ethics Guide to include transparency principles advocated by proponents of the Data Access and Research Transparency (DA-RT) initiative. It gained further momentum in 2014, as a group of editors of leading political science journals issued a joint statement endorsing DA-RT. Subsequently, a growing number of articles have included references to "online appendices," "replication files," "supplements," "online annotations," and similar terms denoting information meant to enable the replication of results.

Today, the basic expectation from authors is to publish any dataset used for quantitative analysis and any additional material necessary for replication, such as the software code used in the analysis or codebook explaining how variables were coded. Moreover, qualitative and mixed-methods researchers are increasingly encouraged to explain why and how they chose and collected their data; to spell out the analytical path through which they arrived at their conclusions; to articulate how every cited piece of primary or secondary evidence is relevant to their claims; and, when possible, to grant access to these sources (Elman, Kapiszewski, and Lupia 2018). One method of documenting the research process is "active citation"—hyperlinking specific arguments to the sources that substantiate them or to excerpts from these sources (Moravcsik 2014). Another technique often mentioned in tandem with active citation is annotating—attaching to specific segments in the main article digital notes that contain additional information on the highlighted text, reminiscent of the comments in a PDF document (Karcher and Weber 2019).

The information era has considerably facilitated replication. In the past, large repositories of raw data were difficult to digitize and expensive to store (Meier 1995, 459). However, recent technological developments have dismantled these obstacles. Thanks to the steep drop in the cost of digital storage and the emergence of numerous academic and non-academic data archives, researchers and journals can now host and share raw and structured replication data freely or at an affordable cost (Alvarez, Key, and Núñez 2018, 422. See also OAA #4).

Sharing quantitative replication data has indeed become a norm in top-tier political science journals. In 2016, Ellen M. Key examined the DA-RT practices of six leading disciplinary journals,[2] concluding that over one-third of the articles in her sample "did not have publicly available replication materials" (Key 2016, 271). We used Key's article as a baseline, checking the 97 articles that were published in the Fall 2019 issues of these journals. Table 1 displays a summary of the findings. As we document in the online appendix, the only articles without appendices are five political theory studies that do not draw on empirical data (refer to the online appendix for our data, methodology, and results). It follows that virtually every evidence-based article published today in leading political science journals offers some type of appendix. When the appendix includes a replication dataset (N=79), the latter includes both data and code in a majority of cases (76).

Unfortunately, offering replication data is not synonymous with being replicable. For example, the inclusion of data does not ensure that researchers will know how to use them. Many appendices are highly compartmentalized—seventy-four articles in our analysis include both a text document and a replication data repository. Out of these seventy-four articles, only six include a hyperlinked reference to the latter. In all other cases, one would have to find, visit, and download files from at least three different web addresses—the main article, the text document, and the replication data repository—to replicate the analysis.

Navigating the different parts of the appendix can be equally daunting. Out of eighty-six text documents in our analysis, only thirty-three have a table of contents that refers to specific page numbers; in eight cases, there are no page numbers at all. Out of seventy-nine replication datasets, only forty-seven include a README file listing the contents of the data repository or a codebook explaining how the variables were coded. Many of these documents are concise, laden with technical jargon (OAA #5), and lacking information on decisions and dilemmas regarding data collection, coding, and analysis, which is crucial for replication. Opening the replication files often requires specific software packages that are not freely downloadable or widely used (see Alvarez, Key, and Núñez 2018, 423).

https://doi.org/10.1017/S1537592720001206 Published online by Cambridge University Press

Table 1.
The accessibility of online appendices in the fall 2019 issues of six political science journals

Journal	No. of Articles	Articles with Appendix	Appendix Includes Text Document	Text Document Includes Table of Contents	Appendix Includes Data Archive	Archive Includes Instructions	Archive Includes Data	Archive Includes Code	Appendix Includes Text Document and Data Archive	Text Document Includes Link to Data Archive
American Journal of Political Science	15	14	14	10	12	12	12	12	12	0
American Political Science Review	14	13	12	4	10	5	10	10	10	1
British Journal of Political Science	17	16	15	1	16	5	16	14	15	2
International Organization	7	7	7	3	2	1	1	1	2	0
Political Analysis	13	13	10	4	13	13	13	13	10	3
The Journal of Politics	31	29	28	11	26	11	26	26	25	0
Total	**97**	**92**	**86**	**33**	**79**	**47**	**78**	**76**	**74**	**6**

Note: Dataset, codebook, and additional tables can be found in the online appendix of this article and at https://github.com/jonathan-grossman/Appendices/blob/master/Codebook.pdf.

Thus, on the one hand, most current online appendices fulfill formal reproducibility requirements. They are meant to be read by a fraction of those who read the article—a narrow group of scholars who share the authors' areas of expertise, methodological propensities, and computer skills, and are therefore capable of using the information in the appendix to replicate the study's findings. This access is a praiseworthy achievement of the replication movement, whose importance cannot be overemphasized. On the other hand, the field must establish best practices to work toward full reproducibility. Appendices could become relevant for a much larger segment of researchers, who might benefit from the information in them but are unlikely to invest the time and resources necessary to decipher them. For this to happen, we must change the way we think about appendices.

Replication and Beyond: Why We Need (and Deserve) Accessible Appendices

Making our scholarship clearer does not mean that we need to "dumb down" our argument and methodological explanations or abandon our rigor; scholarly quality and accessibility need not be mutually exclusive (Flinders 2015, 75; OAA #6). Rather, we should strive to create appendices that are both comprehensive and clear, such that specialists in our field can easily understand and navigate them when attempting to replicate our research.

Moreover, while meeting the replication standard is the *raison d'être* of online appendices, their potential contribution to scholarship extends beyond reproducibility. Appendices are unlimited in length and, ideally, more potent and flexible than published articles in terms of data visualization, interactivity, and the use of audiovisual materials. These traits render appendices a powerful medium through which to make strong, convincing claims that cannot be made in printed articles, either to adhere to formal traditions of academic publishing or because they require extratextual media. In this regard, we can learn from practitioners in the burgeoning field of data journalism who create beautifully animated or interactive visualizations of their arguments (Hahn and Stalph 2018). While such data storytelling loses its impact when converted to traditional textual forms, digital appendices are an ideal medium for presenting it.

In addition, online appendices can enhance the "understandability and persuasiveness" of our research (Lupia and Elman 2014, 22). As Paul Musgrave and Sebastian Karcher state regarding online annotations, the unique features of appendices can help us to "showcase the depth" of our research and "make it easier for readers to go beyond the article" (Musgrave and Karcher 2018; OAA #7). Appendices make an excellent platform over which to provide our readers with historical, cultural, and social context, which is indispensable in evidence-based research, including quantitatively oriented works

https://doi.org/10.1017/S1537592720001206 Published online by Cambridge University Press

(Lieberman 2010). For example, appendices can offer a bibliographic section with hyperlinked references to sources (or, when applicable, full texts) that are not cited in the paper yet give additional insight into the subject in question; a background section more elaborate than the one in the main text; or structured timelines and chronologies of pertinent events, which are often essential in historically oriented analysis (see for example the appendix of Ricks and Liu 2018). Specific words and sentences in the appendix can include hyperlinks to such resources as Wikipedia entries or open access articles for key terms, actors, and events (see Dunleavy, Park, and Taylor 2018 for examples; OAA #8).

Appendices are also an opportunity for more dynamic research. Based on traditional academic journals, a printed article is finalized upon publication. To correct errors, make comments, or add variables and observations to the analysis, one must publish another piece—a comment, follow-up, erratum, letter to the editor, blog post, and so forth. A dynamic online appendix would enable the author or journal to incorporate such contributions in a transparent and timestamped way that tracks and presents changes to the appendix while preserving the original text of the article.

How to Create Better Online Appendices

Political science is a pluralistic discipline with diverse research traditions that require flexibility in preparing appendices. The types of data to be included in appendices vary from photos of material objects and ancient manuscripts, to reflexive field notes and interview recordings, to software syntax and output (OAA #9). While this diversity means that appendices emanating from different research approaches might look very different, their core purposes are the same: every appendix should allow editors, reviewers, and readers to trace the authors' steps, replicate their work, and delve into additional helpful or interesting information. To attain these goals, both authors and journals need to take action (see Ishiyama 2014; OAA #10). Authors should create appendices to be intuitive and integrated, while journals should review appendices to ensure that they meet the basic requirements as well as provide a suitable platform to host them. In this section, we propose principles for appendix development followed by best practices for journals and authors to implement these principles.

Striving for *intuitiveness* means that appendices should be reader-centered rather than author-centered. A political scientist, even without familiarity with the author's methodology or field of expertise, should be able to navigate the appendix, access the files, and understand their nature, purpose, and the type of software needed to open them (OAA #11). We should attempt to use common file formats that can be opened with freely available software —for example, PDF, TXT, or RTF file extensions for text documents or CSV files for structured data. Files that require specialized software should be uploaded alongside, and not in place of, more common file types (OAA #12). Readers should be able to access single files instead of having to download the entire repository (which is often compressed into a very large file). Text documents should have tables of contents, page numbers, and section headings; figures and tables should be embedded in them rather than uploaded separately.

Integrity calls for authors to treat all parts of their appendix, as well as the main article, as components of a single, comprehensive body of knowledge. The present situation, in which "supplementary information" and "replication data" are two distinct entities detached from each other, is untenable in terms of replicability. In order for this body to function, its organs must seamlessly communicate with one another through references and hyperlinks. To ensure integrity, every appendix should have an "anchor"—a web page or text document serving as the roadmap for the entire appendix. Like the table of contents in a text document or the README file in a data repository, this page should utilize clear language and link to every section or file in the appendix (Alvarez, Key, and Núñez 2018, 424; OAA #13).

To implement principles of integrity and intuitiveness in appendices, every journal needs to adopt some *standards* regarding the way authors create, store, display, and format their appendices (OAA #14). This may seem counterintuitive; after all, appendices are supposed to allow authors to break free from the cage of rigid and sometimes arbitrary formalities that journals impose upon authors. However, this freedom must not come at the expense of accessibility and comprehensiveness. Because of the eclectic and idiosyncratic design of current online appendices (which stands in contrast to the relatively regulated structure of articles), understanding the internal logic of each appendix becomes an onerous task that few researchers would take upon themselves.

To resolve this problem, journal submission guidelines should clarify what must be included in the appendix in order for the article to be publishable, in accordance with the prevalent norms in the journal's respective field and with the type of data and research methods in the article. For example, the appendix of an article that features the quantitative analysis of structured data should consist of a detailed roadmap for every stage in the research, the data that are being analyzed, the codebook, and an analytical script—the code or syntax of the statistical analysis performed. Authors of qualitative appendices should describe their selection of sources, methodological and analytical choices, and hypotheses tested (including ones that were rejected). They should also aim at exposing readers to the raw data—documents, recorded interviews, and other sources or parts of them—that led to their conclusions, insofar as these

https://doi.org/10.1017/S1537592720001206 Published online by Cambridge University Press

materials can be legally and ethically shared (see Moravcsik 2014, 686; Tripp 2018; OAA #15). While some journals have already developed rather detailed instructions for creating appendices, these standards need to be better enforced and more widely implemented (OAA #16).

Finally, journal editors should address the *platform* question, that is, where and how to host the appendix. The most suitable outlets are journal websites, content management systems, and version control systems. All these platforms entail a certain tradeoff, which we proceed to identify for each one.

Journal websites. The websites that journals use for publishing online articles are a natural fit for appendices. The interface of these websites can be intuitive and easy to follow. However, to create and sustain such appendices, publishers and editors would have to allocate substantial resources for designated experts responsible for reviewing the materials that authors provide, reformatting them in accordance with each journal's style, ensuring their integrity by linking each part of the article to the corresponding part of the appendix, uploading them, and maintaining them. Given the high costs of such an endeavor, journals may be unable to take on the maintenance of appendices themselves (OAA #17).

Content management systems (CMS). These include blogs, wikis, and platforms such as Wordpress or Google Sites. A CMS provides authors with intuitive and flexible graphical user interface environments for creating dynamic appendices and for modifying their content if needed. Nonetheless, journals that use a CMS, especially those that are not hosted on their publishers' servers, delegate the responsibility for appendix maintenance to the authors. The editors can verify the existence of an appendix as well as its adherence to the journal's requirements only prior to publication. Once an article is published, it would be almost impossible to enforce authors to keep appendices accessible and updated (#OAA 18).

Version control systems (VCS). The neuroscientist and research scientist Patrick Beukema recently proposed to address replication problems in science by borrowing core principles from the software industry. With respect to political science, this application is not as far-fetched as it may appear at first glance: the political process, according to Harold Lasswell (1950), determines "who gets what, when, and how," while version control can help researchers establish "who did what, and when" (Beukema 2018).

VCSs such as GitHub or the Open Science Framework are dynamic worklogs. They record or take a "snapshot" of a project whenever it changes and allow comparison across time. Every action—such as uploading a new file or changing an existing line of code—prompts the creator to add a textual explanation that is recorded and time-stamped. Moreover, VCSs are highly dynamic tools aimed at identifying and correcting coding errors; as such, they provide a transparent outlet for replication and amendments: scholars can download data, correct mistakes in them, add new variables or observations, and propose to the author to merge the modified dataset into the existing one. To exemplify the potential of a VCS, we offer the appendix of this article (which includes online annotations, methodological explanations, and replication materials) as both a GitHub appendix and a traditional PDF document.[3] However, despite these advantages, at this juncture, VCSs are mostly used by teams of experienced coders. Far from intuitive, they are likely to frustrate many authors and readers who lack technical skills.

Recognizing these benefits and disadvantages, journals must choose or seek out platforms that work toward greater intuitiveness for their readers, contributors, and editors.

Conclusion

In a sense, replication in political science has never been better. In the discipline's leading journals, the vast majority of recent articles offer online appendices that include text documents, data files, and software syntax, which in theory should suffice for the replication of the article's analysis. In many cases, however, full replication has yet to become feasible because appendices are disintegrated, unintuitive, and inaccessible.

Having met the replication standard, broadly defined, it is now time for political scientists to create coherent, comprehensible, and integrated online appendices. Such appendices will not only ensure fuller and better replication, but also have the potential to make the article accessible for a larger number of scholars to enhance its overall impact. For this to happen, the disciplinary community must first establish minimal standards of how appendices should look and where they should be stored. No less important, we should incentivize authors, journals, and publishers to meet these standards. We believe that these challenges and their solutions will represent the next step in political science's ongoing quest for transparency and replicability.

Supplementary Materials

To view supplementary material for this article, please visit http://dx.doi.org/10.1017/S1537592720001206.

Notes

1 A dynamic online appendix for this article is available at https://github.com/jonathan-grossman/Appendices.
2 *American Political Science Review*, *American Journal of Political Science*, *British Journal of Political Science*, *International Organization*, *The Journal of Politics*, and *Political Analysis.*
3 https://github.com/jonathan-grossman/Appendices.

References

Alvarez, R. Michael, Ellen M. Key, and Lucas Núñez. 2018. "Research Replication: Practical Considerations." *PS: Political Science & Politics* 51(2): 422–26.

Beukema, Patrick. 2018. "Modernizing Academic Data Science." *Towards Data Science.* Retrieved March 6, 2020 (https://towardsdatascience.com/modernizing-academic-data-science-4213794a278e).

Dunleavy, Patrick, Alice Park, and Ros Taylor, eds. 2018. *The UK's Changing Democracy: The 2018 Democratic Audit*. London: LSE Press. Retrieved March 6, 2020 (https://www.democraticaudit.com/the-uks-changing-democracy-the-2018-democratic-audit/).

Elman, Colin, Diana Kapiszewski, and Arthur Lupia. 2018. "Transparent Social Inquiry: Implications for Political Science." *Annual Review of Political Science* 21 (1): 29–47.

Flinders, Matthew. 2015. "The Rediscovery of the Political Imagination." In *The Relevance of Political Science*, ed. Gerry Stoker, B. Guy Peters, and Jon Pierre, 65–83. Basingstoke: Palgrave Macmillan.

Gibson, James L. 1995. "Cautious Reflections on a Data-Archiving Policy for Political Science." *PS: Political Science & Politics* 28(3): 473–76.

Golder, Matt, and Sona N. Golder. 2016. "Symposium: Data Access and Research Transparency (DA-RT)." *Comparative Politics Newsletter* 26(1): 10–64.

Hahn, Oliver, and Florian Stalph, eds. 2018. *Digital Investigative Journalism: Data, Visual Analytics and Innovative Methodologies in International Reporting*. Cham: Palgrave Macmillan.

Ishiyama, John. 2014. "Replication, Research Transparency, and Journal Publications: Individualism, Community Models, and the Future of Replication Studies." *PS: Political Science & Politics* 47 (1): 78–83.

Karcher, Sebastian, and Nicholas Weber. 2019. "Annotation for Transparent Inquiry: Transparent Data and Analysis for Qualitative Research." *IASSIST Quarterly* 43(2). Retrieved March 6, 2020 (https://doi.org/10.29173/iq959).

Key, Ellen M. 2016. "How Are We Doing? Data Access and Replication in Political Science." *PS: Political Science & Politics* 49(2): 268–72.

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

Lasswell, Harold Dwight. 1950. *Politics: Who Gets What, When, How*. New York: P. Smith.

Lieberman, Evan S. 2010. "Bridging the Qualitative-Quantitative Divide: Best Practices in the Development of Historically Oriented Replication Databases." *Annual Review of Political Science* 13(1): 37–59.

Lupia, Arthur, and Colin Elman. 2014. "Openness in Political Science: Data Access and Research Transparency: Introduction." *PS: Political Science & Politics* 47(1): 19–42.

Meier, Kenneth J. 1995. "Replication: A View from the Streets." *PS: Political Science & Politics* 28(3): 456–59.

Moravcsik, Andrew. 2014. "Trust, but Verify: The Transparency Revolution and Qualitative International Relations." *Security Studies* 23(4): 663–88.

Musgrave, Paul, and Sebastian Karcher. 2018. "Seven Reasons We Use Annotation for Transparent Inquiry (ATI)." *Duck of Minerva.* Retrieved March 27, 2020 (https://duckofminerva.com/2018/05/seven-reasons-we-use-annotation-for-transparent-inquiry-ati.html).

Ricks, Jacob I., and Amy H. Liu. 2018. "Process-Tracing Research Designs: A Practical Guide." *PS: Political Science & Politics* 51(4): 842–46.

Tripp, Aili Mari. 2018. "Transparency and Integrity in Conducting Field Research on Politics in Challenging Contexts." *Perspectives on Politics* 16(3): 728–38.

https://doi.org/10.1017/S1537592720001206 Published online by Cambridge University Press