

Original Article

Qualitative and Multi-Method Research

Fall 2024, Volume 22.2

<https://doi.org/10.5281/zenodo.14062851>

Building Historically Oriented Datasets: A Practical Guide

Erica De Bruin
Hamilton College

Kristen A. Harkness
University of St. Andrews

Jun Koga Sudduth
Indiana University Bloomington

Introduction

Political science has been experiencing an upsurge in the construction of original, historically-oriented datasets. Whether cross-national or subnational, these data collection efforts are increasingly recognized as invaluable to the field—providing resources of enormous long-term benefit and enabling richer empirical analysis of a broader range of research questions.¹ There is a large and rich literature on conceptualization and measurement in the social sciences, which provides an important foundation for such dataset construction (e.g., Adcock and Collier 2001; Goertz 2006; Sartori 2009; Schedler 2012). There is also work that identifies best practices or principles scholars should follow in building

historically oriented datasets—those that involve “the translation of narrative records of events and processes into numerical scores in the form of indexes, scales, or dummy variables indicating the presence or absence of some trait, such as regime type or degree of conflict” (Lieberman 2010, 39; Salehyan 2015).² At a pragmatic level, however, there are few resources on *how* to collect, code, and document historically oriented data.

To aid scholars embarking on efforts to build such historical datasets, this article presents a practical guide for dataset construction. It is geared towards graduate students, early career scholars, and others without extensive financial resources or a large research team. It may also be useful to scholars piloting data collection processes in order to secure funding for subsequent scaling-up. Our goal is to make transparent the choices and trade-offs involved during the data generation process, while providing helpful tips and highlighting pitfalls to avoid.

We provide a sequenced approach for historical data construction. We first discuss how to limit the scope of data collection to make it feasible for researchers with time and budgetary constraints. We then discuss sources and the trickiness of coding variables from collated narrative

1 The limitations and biases of such data are also quite well-known (see e.g., Bagozzi et al. 2019; Dietrich and Eck 2020; Gohdes and Price 2013; Hug 2003; Weidmann 2015).

2 Lieberman (2010), for example, emphasizes the principles of proximity of observations, transparency in citations, certainty of the historical record, and attention to valid comparison.

information, advocating pilot projects and a flexible, iterative process of refinement. Finally, we identify principles that should guide coding documentation and describe how to manage what can be an overwhelming amount of material by creating good workflow processes. Throughout, we draw upon our own experiences, failures, and hard-won lessons learned in constructing cross-national time-series datasets.³

Determining the Scope of Your Dataset

Embarking on historical data collection is daunting. Political processes often have no clean beginnings, neat endings, or clearly defined boundaries. Ideally, many of our projects would code hundreds of variables far back into the mists of time, include every conceivable unit of observation, and disaggregate as much as possible. But that task would never end. Thinking sensibly about your own time is important, especially for early career scholars. Job market, tenure clock, and promotion processes create extraordinary pressures to publish quickly. But only the fortunate few have sufficient institutional resources or grant funding to hire large teams of research assistants to gather and code data for a dissertation or first book.

Data collection becomes far more manageable if you can limit its scope. Here, the tricky part is balancing such reductions against the validity and richness of your empirical analyses—you want to ensure that your scoping decisions will still ensure you have leverage over the type of variation that is most relevant for answering your research questions and generating or testing your theories.

Scoping decisions are also fundamentally intertwined with operationalization—the art of turning theoretical concepts into observable and measurable indicators. Operationalization is well-covered in the existing literature, and we would strongly recommend a thorough read (see, for example, Adcock and Collier 2001; Collier, LaPorte, and Seawright 2012; Goertz 2006; Pepinsky 2007;

Sartori, 1970). From a practical standpoint, we would also emphasize, before developing your own coding guidelines, first looking at how other scholars have previously operationalized your concepts and closely related ones. This allows you to understand what work has already been done, standard practices in the field, and how other scholars have thought through tricky issues.

In our collective experiences, we have used four strategies to limit the scope of data collection: (1) winnowing variables; (2) limiting temporal or geographical coverage; (3) randomly sampling cases; and (4) reducing granularity. There are both benefits and risks to each.

Winnowing variables. Conceptualizing, measuring, gathering information about, and coding a small handful of variables—as opposed to dozens or hundreds—is intuitively less time-consuming. Such laser focus, however, depends heavily on well-developed theory. You must first identify a narrow set of independent and dependent variables that will enable compelling empirical tests. This includes thinking through how to potentially test causal mechanisms and/or alternative explanations. The greatest risk of this strategy is narrowing too early. It is far easier, from the outset, to track multiple types of information from a set of sources than to realize later that you must add another key variable and need to return to your source material to do so. Pilot projects can help thread this needle, as discussed later, allowing the researcher to start broader and then clarify which variables may be too difficult or time consuming to code.

Limiting temporal or geographic coverage. Every project makes decisions over time and space. Strong theory, and careful consideration of variation, not only strengthen research design but can conserve resources as well. What type of variation will give you the most leverage over your research questions: across cases, or within cases, over time? To the extent that you want to leverage temporal variation, do you need annual, monthly, or daily data? Or would snapshots at specific historical

3 These include datasets on military purges, ethnic stacking in Africa, and features of state security and police forces (e.g., De Bruin 2021, 2022; Harkness 2022; Sudduth 2021).

moments suffice? Is there a narrower time period that would make for an ideal test of the theory or capture the most important variation?

Keep in mind that you can expand coverage after you have demonstrated the value of your data. In her book on building and dismantling ethnic armies, for example, Kristen focused her data collection efforts on the immediate post-colonial period and the third wave of democratization (Harkness 2018). Only later—with more time and more funding—did she build a comprehensive cross-national time-series dataset (Harkness 2022). Similarly, for her pilot study of civilian elite purges, Jun limited her initial data collection to the years 1985-1996. This allowed investigation of possible systematic changes in purge trends since the end of the Cold War, while reducing data collection time.

Also worth considering is the nature of real-world variation. For instance, if there is little temporal variation, the value of collecting time-series data is much less clear. Erica found that coding features of police forces in the first and last years of counterinsurgency campaigns took a fraction of the time than coding annually without much loss in data richness, as very few police forces underwent significant organizational change in between (De Bruin 2022). And it may not be worth coding data before or after the temporal ranges of other datasets containing variables you—or the users of your data—may frequently use.

Similarly, narrowing to a specific geographic region or sub-region can make data collection more manageable. This was the strategy Kristen used in her dataset on ethnic stacking; the dataset focuses on Africa, where the recruitment and promotion of coethnics has been a crucial component of rulers' efforts to control their militaries (Harkness 2022). It has the added benefits of limiting conceptual stretching and honing the researcher's contextual knowledge in a beneficial way. Deep regional or country expertise can also generate public engagement and impact opportunities. Yet, a narrow geographic focus has important implications for generalizability. Other scholars

may perceive your theory as only relevant to the region where it was empirically tested. Your data will also have diminished utility to other scholars who either focus their work on a different region or want to conduct global analyses. Finally, some journals discourage regional datasets, limiting—but by no means eliminating—publication outlets.⁴

Randomly sampling cases. Another strategy is to only collect data on a manageable subset of randomly selected cases, as Erica did with the State Security Forces Dataset (De Bruin 2021). One could choose every sixth country listed alphabetically or use a random number generator to select observations. With this approach, geographic and temporal limits are minimized, allowing for better integration with existing (especially cross-national) datasets, and inferences can still be drawn to the wider population. But there are also drawbacks. A random sample may end up excluding important cases within the literature to which your project speaks. It could also require you to develop contextual expertise on perhaps an uncomfortably broad set of cases. Finally, concerns over missing observations may still compel future studies to draw their variables from other datasets—regardless of improvements in reliability or validity that your dataset might be able to make.

Reducing granularity. Finally, coding simple binary variables or categorical variables with a small number of outcomes is usually less time consuming than coding continuous variables. It also enables wider coverage and higher recovery rates, minimizing missing data. For example, Jun found it easier to code whether a leader purged officers from the military at all (0 or 1) than to measure consistently and reliably how many officers were purged (a continuous variable) (Sudduth 2017). Some granularity could be preserved, without too much additional work, by creating categories for different levels of purges (e.g., none, less than ten, 11 to 100, more than 100) (Sudduth 2021). Again, it is worth thinking hard about whether fine-grained distinctions between categories will map onto your theoretical constructs in a meaningful way,

4 For example, the *Journal of Peace Research* has historically tended to publish datasets with global coverage, as well as subnational datasets; others, such as *Conflict Management and Peace Science*, have welcomed those with a regional focus.

or whether you can do without them. The trade-off is information loss —missing details that could be important for later empirical tests and for ensuring other researchers, testing different theories and causal mechanisms, find the data useful.

Regardless of how you ultimately balance this trade-off, we recommend capturing the most granular information that you can find in your source material at the beginning of a project. It is much easier to code at higher levels of abstraction, given a rich information base, than to go back and re-collect more detailed information later in the project. Furthermore, more detailed information can help you understand processes and dynamics better, as well as provide useful anecdotes or serve as the basis for qualitative case studies.

Identifying and Selecting Source Material

Once you know the general scope of the data you need to collect, the next step is to think about sources. Rather than delve immediately into primary sources, we have found it valuable to first gather what information you can from existing academic scholarship (particularly country-specific historiography and ethnography), historical dictionaries, annual surveys, and other such resources. Then you can strategically think through the gaps that need filling with more time-intensive primary sources.

It is important to remember that all sources are subject to data generating processes that create biases in what information is recorded and preserved. News sources are easy to access as most academic institutions subscribe to searchable online databases, including LexisNexis and Keesings, which provide broad global coverage, as well as national and local newspapers. However, they are shaped by audience appeal and may be biased in favor of particular countries, urban areas, and the

most “news-worthy” or visible events (Croicu and Eck 2022; Dietrich and Eck 2020; Parkinson 2024; Weidmann 2015). Historical archives can provide astoundingly rich and unique information, enabling greater disaggregation and more direct capture of key concepts. However, access to archives can be tricky and, at a minimum, requires the time and funding to travel to the location of preservation.⁵ Archives also contain their own biases, which reflect the incentives of actors recording and releasing the records (Balcells and Sullivan 2018; Lee 2022), issues likely exacerbated by the (often partial) digitization of archival collections (Kim 2022). Whatever the sources you use, we would thus echo Salehyan’s (2015) encouragement to think hard about what may be missing from them, and to triangulate contested information from multiple sources wherever possible.⁶

Using Pilot Projects Effectively

Once you have developed initial ideas of how you want to measure your concepts (e.g., coding guidelines) and where you will find relevant historical information (sources), we strongly recommend embarking on a pilot coding project.⁷ Pilot projects enable you to iteratively refine your processes of data collection and coding and reflect on measurement validity: do the numerical scores produced map back onto the original concepts and theoretical constructs in a meaningful way? Could revisions to coding practices improve such validity? Carefully crafted pilot projects generate confidence in your approach before sinking extensive time and resources into a particular coding scheme.

We have found it helpful to strategically select a small sample of observations to code that will “stress test” your initial coding rules by making you consider the full range of outcomes on your variables, highlight gray zones between coding categories, and probe the limits of source material.

⁵ For a thoughtful discussion of the politics of archives, and other ethical issues in archival research on political violence, see Subotić (2021).

⁶ Both Lee (2022) and Kim (2022) provide excellent guides to how to anticipate and help mitigate biases in different forms of archival research.

⁷ Budget allowing, it may be helpful to hire one or two research assistants at this stage to help you pilot. You can provide them the guidelines and the same set of sample cases to check the extent to which their codings converge.

Sample cases from different geographical regions or sub-regions and from different historical eras, decades, or years. Pick some difficult cases that do not easily fit your concepts and chosen measures. Include some observations with a wealth of historical materials—an overabundance of sources that will generate differing and sometimes contradictory interpretations. But also include observations with a dearth of sources that will stretch your capacity to find enough material and help highlight which variables it might not be feasible to collect data on across your cases.

Pilot projects usually reveal useful time-saving lessons such as coding rules that need modification and variables that cannot be systematically coded across cases. Keep a flexible mindset and remind yourself that this is the whole point of the pilot—to find the problems and improve measurement validity by iteratively modifying and re-test your approach. You can re-pilot as many times as needed. Then, when you feel confident in your processes and coding guidelines, scale up and tackle the rest of the data collection project.

A final tip: Keep extensive notes for yourself during your pilot. Track which variables were most challenging to find sufficient information on; where you had trouble coding the information you found; sources that were particularly useful (or turned out to be irrelevant); and where interesting and important features of the cases were not yet captured by your coding scheme. Your future self will be very grateful for detailed tracking of your own thoughts and decisions over tricky issues and difficult cases.

Documenting Your Coding Decisions

In documenting your coding decisions, the aim is to ensure transparency and, to the extent possible, replicability. Transparency centers on creating a trail of documentation that allows others to understand each step of how you built the data, the decisions involved, and the sources consulted.⁸ It gives meaning to the numbers, enabling good inferences.

Replicability sets an even higher bar. Some would argue that after reading your full documentation, other scholars should be able to reach the same coding decisions as you did, in the vast majority of most cases. This is a high bar indeed, and one that may overlook the unavoidable messiness of social data.

There will always be debates over data subjectivity, as human judgment is fundamental to translating often incomplete, complex, and sometimes contradictory narrative records into numerical scores. Take the recent debate over whether and to what extent democracy is backsliding globally. Little and Meng (2024, 151) highlight a discrepancy between “objective” indicators of democracy, which they describe as “based in fact,” and “subjective” indicators, which depend on a combination of fact and coder judgement. However, even seemingly objective indicators often require multiple judgments by human coders (Schedler 2012). Knutsen et al. (2024) provide the example of coding parliamentary vote share—while theoretically fact-based, coders must still make choices about how to code independents, which electoral rounds to consider, and other issues.

Limitations in source material may also produce ambiguous or conflicting information that must be adjudicated. Jun found this to be the case in coding military purges, which she defined as leaders’ actions to dismiss, demote, expel, arrest, or kill individuals within their security apparatus. In some instances, there was evidence that top-ranking officers left their positions unexpectedly, but whether they resigned voluntarily or were forced out remained unclear (Sudduth 2021).

More generally, many of the concepts most integral to the study of politics are deeply normatively important, and thus essentially contested (Gallie 1955-56). As a result, the expectation should not be that our codings will be accepted uncritically. Indeed, debates over coding can be instrumental in advancing our understanding of complex political phenomena. The debate over whether the January 6

⁸ Many political science journals now subscribe to Data Access and Research Transparency (DA-RT) principles, which aim to increase transparency in social science (<https://www.dartstatement.org/>).

attack on the U.S. Capitol counted as a coup d'état, example, drew attention to a growing divergence between academic and journalistic uses of the term; the extent to which major coup datasets disagreed about specific cases (Chin, Carter, and Wight 2021); and the need to better map the conceptual terrain of other types of antidemocratic actions, such as “self-coups,” which share some features with coups but also diverge from them in important ways (Powell et al. 2022). What good coding and documentation practices do allow us to do is minimize unnecessary disagreement by making explicit the choices we have made and enabling other scholars to evaluate for themselves whether they would come to the same decisions about our cases.

Managing Documentation and Workflow

Handling all the source materials, citation information, qualitative narratives, coding guidelines, and spreadsheets involved in these projects, just for one's own personal use, is a huge task. We cannot stress enough how vital it is to develop good systems, right from the beginning, to track and organize your project. Fixing mistakes can be costly and time consuming. When copy-editing her book, for instance, Kristen had to return to archives she visited as an inexperienced graduate student to fix three footnotes that were missing some small but essential piece of citation information (e.g., part of a document title, the recipient of a letter).

While individualized, the systems we have each developed to manage documentation and workflow revolve around four types of documents that we combine and layer: (1) coding guidelines; (2) source notes; (3) case or coding notes; and (4) data spreadsheets.

First, coding guidelines (the “codebook”) contain the final rules for data coding. They should be ultimately published alongside the dataset. For each variable, they should describe the underlying

concept and how it has been measured. The best coding guidelines also discuss why you made certain decisions and what reasonable alternatives were considered (but ultimately dismissed). Your thinking will change over the course of the project given the iterative process of flexibility and refinement that we suggest imbues all historic data creation. This makes it crucial to keep “in progress” notes and preserve any iteration of coding guidelines used to produce coded data (linked through version numbers to the associated spreadsheets). Draft coding guidelines can also be shared with advisors, mentors, and colleagues for feedback before the final, cleaned-up version is publicly released.

Second, *source notes* are an important intermediary step between learning relevant information and using it to code your data. This step is tempting to skip. However, for data construction that relies on primary sources or requires more interpretation, it is very useful to preserve nuggets of information in the broader contexts in which they were found. Source notes are especially useful for tracking information from archival documents, interviews, ethnography, newspaper articles, and social media posts.⁹ Source notes are also a place to record overly detailed citations, interesting tidbits whose relevance is not immediately apparent, and one's own evolving thinking. The supplementary material included online provides an example of Kristen's source notes on an archive dossier.

Third, *case or coding notes* compile information across multiple sources for a particular case or observation. They can be in any format—from polished narratives to bullet points to a series of questions and answers—but should be rationally organized around the original variables that need coding. The idea is to collate, for that observation, all the various relevant pieces of information from the underlying sources. Citation tracking is critical, and we recommend refraining from any clustering: cite at the sentence level and keep sources distinct even where they contain the same basic fact.

⁹ Ethical and copyright considerations permitting, you may also consider publicly releasing primary source images, notes, or links along with the dataset.

We also endorse tracking “consulted” sources, including those that you do not draw from, so you avoid rereading them. Erica learned this the hard way when she came across her own hand-written post-it note in a book which, it turned out, she had requested via InterLibrary Loan and consulted a year earlier.

Developing a template for case study notes helps guide data collection (and is vital for employing research assistants). Questions to consider when designing a template include: What background information would help understand the context of this case? What do you need to know to code each variable? Is there other information you want to track, even if it cannot be compiled systematically? The extra material often provides a rich resource for later publications, including anecdotally illustrating theoretical points or fleshing out qualitative case studies. For maximum transparency, case study notes can be published alongside the dataset—either in template form or after rewriting into more polished narratives. This renders visible the underpinning information and justifications for why variables obtained the codings they did. The supplementary material includes three examples of different ways to structure case study notes from Erica and Jun.

Finally, *data spreadsheets* record the numeric or categorical codings assigned to each observation. In their final form—you may go through several pilot or interim versions—they provide you and other potential users with the actual quantitative data. You might produce one large spreadsheet, or several data tables linked together through identifiers. For example, datasets are often published in different versions to distinguish between related units of analysis (e.g., conflict actor and conflict-spell, leader-spell and time-series, or event and time-series).

There are many valuable, published resources on good data management practices (see, for example, Weidmann 2023). Just a few pointers to keep in mind from our experience: Use existing coding conventions for identifying countries, conflict actors, etc., to make merging your dataset with others easier. Try to ensure variable names have enough content for you and future users to understand and remember what they capture

without constantly referring to the codebook (e.g., v1, v2, v3 are not that helpful, nor are self-invented acronyms). Do some good spot-checking and quality control at the end—it is easy to make errors while inputting loads of data into a spreadsheet. You can also set up some automated quality control checks within your spreadsheets (e.g., if the variable is supposed to be 1-5, and there's a value of 10, you know it's an error). Finally, generate a comma-separated values (CSV) file for sharing purposes that is easy to use with any software program.

Sharing and Updating Your Data

When you are ready to share your data, posting it in an online data repository like the Harvard Dataverse (<https://dataverse.harvard.edu/>) increases data accessibility beyond your own networks and allows you to track the number of downloads. You can also choose to gather information on who uses your data by requesting names, institutions, and positions from users before they download your datasets. However, keep in mind that requesting this information may create a barrier to access for users uncomfortable with sharing such personal data, so consider how important it is to you.

Finally, while one may be tempted to hang the victory banner at this stage, it is useful to consider how you want to process queries or disputes, as well as how you will revise and release updated versions of your dataset. Periodic updates may be necessary to extend your data collection forward in time. New records may be declassified, or more detailed historiographies published, which change your interpretation of cases. And ideally, lots of other scholars will use and scrutinize your data, which will bring the occasional data entry or coding error to light, even where it does not provoke broader debate over your indicators. As a field, we should think more deeply about how to support periodic cleaning and updating of our data resources—an important task for which little financial support is typically available. In the meantime, developing your own plan for revisions will help ensure your data remains relevant to users.

Conclusions

The construction of original, historical-oriented datasets allows political scientists to document patterns and to develop and test theories about the political world. Our aim, in this article, was to provide a practical guide to building datasets that highlights important tradeoffs and provides advice to assist researchers undertaking such efforts for the first time. We hope that it helps get more data projects off the ground.

Acknowledgements

For helpful feedback, we are grateful to an anonymous reviewer, Carolyn Holmes, Nils Weidmann, Agnes Yu, and other participants in the 2022 American Political Science Association Annual Meeting panel on “Critical Challenges and Solutions Towards (Re)Constructing Conflict Datasets.”

References

- Adcock, Robert, and David Collier. 2001. "Measurement Validity: A Shared Standard for Qualitative and Quantitative Research." *American Political Science Review* 95, no. 3 (September): 529–46.
- Bagozzi, Benjamin E., Patrick T. Brandt, John R. Freeman, Jennifer S. Holmes, Alisha Kim, Agustin Palao Mendizabal, and Carly Potz-Nielsen. 2019. "The Prevalence and Severity of Underreporting Bias in Machine- and Human-Coded Data." *Political Science Research and Methods* 7, no. 3 (July): 641–49. <https://doi.org/10.1017/psrm.2018.11>.
- Balcells, Laia, and Christopher M. Sullivan. 2018. "New Findings from Conflict Archives: An Introduction and Methodological Framework." *Journal of Peace Research* 55, no. 2 (March): 137–46. <https://doi.org/10.1177/0022343317750>.
- Chin, John J., David B. Carter, and Joseph G. Wight. 2021. "The Varieties of Coups D'état: Introducing the Colpus Dataset." *International Studies Quarterly* 65, no. 4 (December): 1040–51.
- Collier, David, Jody LaPorte, and Jason Seawright. 2012. "Putting Typologies to Work: Concept Formation, Measurement, and Analytic Rigor." *Political Research Quarterly* 65, no. 1 (March): 217–32.
- Croicu, Mihai, and Kristine Eck. 2022. "Reporting of Non-Fatal Conflict Events." *International Interactions* 48 (3): 450–70.
- De Bruin, Erica. 2021. "Mapping Coercive Institutions: The State Security Forces Dataset, 1960-2010." *Journal of Peace Research* 58, no. 2 (March): 315–25.
- De Bruin, Erica. 2022. "Policing Insurgency: Are More Militarized Police More Effective?" *Small Wars & Insurgencies* 33(4-5): 742–66.
- Dietrich, Nick, and Kristine Eck. 2020. "Known Unknowns: Media Bias in the Reporting of Political Violence." *International Interactions* 46(6): 1043–60.
- Gallie, W. Bryce. 1955-1956. "Essentially Contested Concepts." *Proceedings of the Aristotelian Society* 56: 167–98.
- Goertz, Gary. 2006. *Social Science Concepts: A User's Guide*. Princeton: Princeton University Press.
- Gohdes, Anita, and Megan Price. 2013. "First Things First: Assessing Data Quality before Model Quality." *Journal of Conflict Resolution* 57, no. 6 (December): 1090–108.
- Harkness, Kristen A. 2018. *When Soldiers Rebel: Ethnic Armies and Political Instability in Africa*. New York: Cambridge University Press.
- Harkness, Kristen A. 2022. "The Ethnic Stacking in Africa Dataset: When Leaders Use Ascriptive Identity to Build Military Loyalty." *Conflict Management and Peace Science* 39, no. 5 (September): 609–32.
- Hug, Simon. 2003. "Selection Bias in Comparative Research: The Case of Incomplete Data Sets." *Political Analysis* 11, no. 3 (Summer): 255–74.
- Kim, Diana S. 2022. "Taming Abundance: Doing Digital Archival Research (As Political Scientists)." *PS: Political Science & Politics* 55, no. 3 (July): 530–8.
- Knutsen, Carl Henrik, Kyle L. Marquardt, Brigitte Seim, Michael Coppedge, Amanda B. Edgell, Juraj Medzihorsky, Daniel Pemstein, Jan Teorell, John Gerring, and Staffan I. Lindberg. 2024. "Conceptual and Measurement Issues in Assessing Democratic Backsliding." *PS: Political Science & Politics* 57, no. 2 (April): 162-77.
- Lee, Alexander. 2022. "The Library of Babel: How (and How Not) to Use Archival Sources in Political Science." *Journal of Historical Political Economy* 2 (3): 499–526. <http://dx.doi.org/10.1561/115.00000038>.
- 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: 35–59.
- Little, Andrew T., and Anne Meng. 2024. "Measuring Democratic Backsliding." *PS: Political Science & Politics* 57, no. 2 (April): 149-61.

- Parkinson, Sarah E. 2024. "Unreported Realities: The Political Economy of Media-Source Data." *American Political Science Review* 118, no. 3 (August): 1527-32.
- Pepinsky, Thomas B. 2007. "How to Code." Political Concepts Working Paper #18, IPSA Committee on Concepts and Methods, <https://www.concepts-methods.org/Files/WorkingPaper/PC%2018%20Pepinsky.pdf>.
- Powell, Jonathan, Salah Ben Hammou, Amy Erica Smith, Lucas Borba, Drew Holland Kinney, and Mwita Chacha, and Erica De Bruin. 2022. "Forum: A Coup at the Capitol? Conceptualizing Coups and Other Anti-Democratic Actions." *International Studies Review* 24, no. 1 (March): viab062. <https://doi.org/10.1093/isr/viab062>.
- Salehyan, Idean. 2015. "Best Practices in the Collection of Conflict Data." *Journal of Peace Research* 52, no. 1 (January): 105-9.
- Sartori, Giovanni. 1970. "Concepts and Misinformation in Comparative Politics." *American Political Science Review* 64, no. 4 (December): 1033-53.
- . 2009. "Guidelines for Concept Analysis." In *Concepts and Methods in Social Science: The Tradition of Giovanni Sartori*, edited by David Collier and John Gerring, 97-150. New York: Routledge.
- Schedler, Andreas. 2012. "Judgment and Measurement in Political Science." *Perspectives on Politics* 10, no. 1 (March): 21-36.
- Subotić, Jelena. 2021. "Ethics of Archival Research on Political Violence." *Journal of Peace Research* 58, no. 3 (May): 342-54.
- Sudduth, Jun Koga. 2017. "Strategic Logic of Elite Purges in Dictatorships." *Comparative Political Studies* 50, no. 3 (November): 1768-801.
- Sudduth, Jun Koga. 2021. "Purging Militaries: Introducing the Military Purges in Dictatorships (MPD) Dataset." *Journal of Peace Research* 58, no. 4 (July): 870-80.
- Weidmann, Nils B. 2015. "On the Accuracy of Media-Based Conflict Event Data." *Journal of Conflict Resolution* 59, no. 6 (September): 1129-49.
- Weidmann, Nils B. 2023. *Data Management for Social Scientists: From Files to Databases*. New York: Cambridge University Press.