How to judge? Qualitative literacy and process tracing studies
Meegdenburg, H., van

of political science are evaluated, it is reasonable to ask whether they include all the relevant evidence or whether "inconvenient" evidence is ignored—whether the accounts presented are not only diverse and palpable, but whether they hold up against other forms of evidence. I would welcome Small and Calarco's suggestions on how to navigate these complicated waters.

References

Auyero, Javier. 2012. *Patients of the State: The Politics of Waiting in Argentina*. Durham: Duke University Press.

Beach, Derek, and Rasmus Brun Pedersen. 2013. *Process-Tracing Methods: Foundations and Guidelines*. 1st edition. Ann Arbor: University of Michigan Press.

Cyr, Jennifer, and Sara Wallace Goodman, eds. Forthcoming. *Doing Good Qualitative Work*. New York: Oxford University Press.

Fairfield, Tasha, and Andrew E. Charman. 2022. *Social Inquiry and Bayesian Inference: Rethinking Qualitative Research*. Strategies for Social Inquiry. Cambridge: Cambridge University Press.

Fu, Diana, and Erica S. Simmons. 2021. "Ethnographic Approaches to Contentious Politics: The What, How, and Why."

Comparative Political Studies 54, no. 10 (September): 1695–721.

Qualitative and Multi-Method Research Spring 2023, Volume 21.1 https://doi.org/10.5281/zenodo.7921415

How to Judge? Qualitative Literacy and Process Tracing Studies

Hilde van Meegdenburg
Leiden University

We all had papers and research proposals rejected invoking standards that, at least to the qualitative researcher, were clearly not applicable: no external validity; difficult to replicate; mere description; selection on the dependent variable(!). We know the list. As Small and Calarco point out, qualitative research's purview is growing but "qualitative literacy" is not keeping pace. That is, the competency of "others" to assess the quality of qualitative research is often insufficient.

Against this background, the authors set out to provide the reader-reviewer with a set of indicators to help recognize craft and assess qualitative evidence competently. In short, how to distinguish good from not-so-good interview-based and ethnographic research? Now, here is a certain irony in being asked to review academic work that is, essentially, about how to review academic work—and unfortunately for me, Small and Calarco's indicators cannot offer guidance. Therefore, in what follows, I do two things. First, I consider the extent to which the authors achieve their primary objective: to provide reviewers, scholars, journalists, and others with a "nonexclusive set of indicators" to evaluate research drawing on in-depth interviews or participant observation. I argue that although their objective is more than meritorious, they, at times, stray in the direction of advising the practicing scholar more than the reader-reviewer. Second, as a practical exercise, I consider how their indicators fare in relation to another approach to qualitative research: process tracing. Here, I find the indicators carry wider than the data collection methods addressed in the book. In fact, I believe *all* qualitative research would benefit from taking them—exposure, cognitive empathy, heterogeneity, palpability, follow-up, and self-awareness—seriously, and, by extension, that the quality of qualitative research can indeed (despite my initial critique) be assessed along these lines.

First, the objective to equip reader-reviewers with the tools to assess the quality of qualitative work is more than meritorious. We can only hope future reader-reviewers will have such a (tentative) list and the book provides a wonderful basis for it. For one, although the authors say the list is "nonexclusive," I had difficulties coming up with indicators to add. Honesty came to mind; the idea that a researcher should be forthcoming about doubts, considerations, contradictions, changes, limitations, and so forth. But honesty is probably covered by heterogeneity, follow-up, and self-awareness. Second, they introduce their indicators in clear terms supported by easy to follow, well-constructed examples slowly building up in complexity. The non-scholarly or scholarly non-qualitative or non-interpretivist reader should really be able to grasp the examples. Where, then, does the critical point I raise in the introduction come from?

My point is this: the assumption seems to be that when we know what good data looks like, we should be able to recognize it. Although the indicators are extensively discussed, also providing generous advice to the practicing scholar, the reader-reviewer is at times less well served. Especially for the "untrained eye," a clear summary or short list of key terms for each indicator would have been helpful. What are the signs of cognitive empathy or self-awareness in a manuscript? How can we see the researcher elicited heterogenous and palpable data, or that they practiced follow-up? I am aware these (sub)indicators are in the book, it is just that they remain implicit, scattered throughout the pages. I have three reasons to raise this point. First, offering a set of indicators is so explicitly the aim of the book and I am afraid the five primary indicators may be too abstract. Second, I would like reader-reviewers, including myself, to have that summary to pin to their (our) walls. Third, as I hope will be clear from what follows, I believe the indicators do, indeed, offer a good basis to begin to assess the quality of qualitative research and it is high time we, as qualitative scholars, start quibbling its details.

Second, inspired by the book but myself neither an ethnographer nor (solely) relying on in-depth interviews, I consider how the indicators fair in my own field of methodological expertise: process tracing (PT). PT is a method relying on in-depth case studies to expose the *causal mechanism(s)* that brought about a given outcome (Beach and Pedersen 2019; Falleti and Lynch 2009; Goertz 2017; Mahoney 2015; Mayntz 2004; Runhardt 2016). Data, and therewith data collection, can be diverse: interviews are frequently employed, but so is archival research, desk research, and so forth. As Beach and Pedersen argued: "evidence can be any type of material that might be left by the workings of our theorized causal mechanism" (2019, 171). Personally, I use PT to study state foreign policy making and combine PT with a focus on narratives (van Meegdenburg 2019; 2023).

Exposure—Although not treated separately, exposure is put forward by the authors as a precondition for high quality data: without adequate time in the field the other criteria would hardly be achievable. For PT, this is no different. Exposure would be understood more broadly as hours spend searching for data—including on-site explorations and interviewing, but also time spent searching (digital) archives for documents, minutes, reports, newspaper articles, you name it—but as with interview-based or ethnographic research, case saturation is important. And the only way to achieve saturation is through adequate exposure. Like ethnographic and interview-based work, PT can be laborious and time consuming.

Cognitive empathy—Reflecting on my own work, cognitive empathy is relevant, but, here too, in a slightly different way. For PT, interviews are possible but not necessary. Yet, when I am working with archival data or, especially, narratives, my aim is often to try and tease out how the main actors "view[ed] the world and themselves—from their perspective" (Small and Calarco, 23). That is, in laying bare a (social) mechanism, actor motivations, believes, perspectives, and emotions, as well as their causal relevance, are often, if not foregrounded, at least playing an important role in the background. Therefore, even when not producing (all) data in a reactive interaction, cognitive empathy should apply to much PT work. To the extent actor perceptions, meanings, and motivations are part of a mechanism, cognitive empathy should be part of our practice.

Heterogeneity—Directly applicable, heterogeneity is important in PT. If someone were to find every step of a hypothesized mechanism perfectly confirmed, "the reader would have reasons to be suspicious" (Small and Calarco, 62). For me, these reflections were timely. I am currently working on a manuscript aiming to expose particular foreign policy narratives and have to find ways to present the heterogeneity that does, most certainly, exist. I found actors and actions can be ambiguous; narratives are always contested; and archival data and documents rarely speak with a single, perfectly aligned voice. As the authors remined me: heterogeneity in the data does not weaken our conclusions; presenting heterogeneity actually heightens the credibility of the research. And this applies to PT, and the evidence we present for our mechanisms, as well.

Palpability—Palpability, or "the extent to which the reported findings are presented concretely rather than abstractly," is also directly applicable. In the case of PT, however, it may have a very specific implication. Palpable data would be data that is explicitly linked to a *particular step or part of the mechanism*. The more abstract argument generally regards the mechanism as a whole, whilst mechanistic evidence should focus on the concrete points, the steps or parts, in between. In my case, the abstract argument regards discursive justification or rationalization as a mechanism, whilst concrete evidence regards specific discursive interventions at specific points in time. In fact, I think PT, and its analytical focus on mechanisms, begs for—and may therefore help elicit—palpable data. Thinking in terms of processes and mechanisms,also when a more formal PT is not the aim, could help qualitative scholars consider what palpable may mean in their case.

Follow-up—Follow-up in PT is what Beach and Pedersen call "iterative" research (2019, 286). PT is often a back and forth between data and theory, between adjusting expectations and gathering more, specific, new, and different data. In my own research, interview statements would alert me to the existence of certain documents, whilst new documents would send me back to talk to people. In fact, over the course of the project

I changed my research focus; I started focusing on narratives only after the data made clear that my initial suspicions were too simplistic, too blunt to describe the much more nuanced processes I was witnessing. Follow-up, thus, is relevant too.

Self-awareness—Lastly, self-awareness may well be the indicator least explicitly discussed in PT. And yet, of course self-awareness also matters when we are not eliciting data through reactive interactions. Who we are, how we think, read, and look, shapes what we see and hear and what we think is significant and worth following up on. In fact, I think self-awareness may well be a key lesson for non-interview heavy or ethnographic process tracers. For me, this includes an awareness of our biases and the behaviors they elicit. The authors discuss *outgroup bias* in relation to heterogeneity, for PT, especially when starting from theory, a *confirmation bias* may play a similar role. I can only say that this point, at least by me, is taken to heart.

Overall, for I really must round this discussion up, I must conclude that PT, and probably *all* qualitative research, would benefit from taking the indicators seriously, and that the quality of qualitative research can indeed be assessed along these lines. To that extent, the indicators provide ample input for reflection, and the book takes a big step towards advancing qualitative literary. This does raise one final thought; the book really begs for a companion: *Qualitative Writing: A Guide to Writing-Up Ethnographic and Interview Research*. Because—as also occasionally shines through in the discussions—the way that research is written affects whether reader-reviewers can recognize the standards that informed the practice.

References

Beach, Derek, and Rasmus B. Pedersen. 2019. *Process-Tracing Methods: Foundations and Guidelines*. Ann Arbor: University of Michigan Press.

Falleti, Tulia G., and Julia F. Lynch. 2009. "Context and Causal Mechanisms in Political Analysis." *Comparative Political Studies* 42, no. 9 (September): 1143–66.

Goertz, Gary. 2017. *Multimethod Research, Causal Mechanisms, and Case Studies*. Princeton: Princeton University Press.

Mahoney, James. 2015. "Process Tracing and Historical Explanation." *Security Studies* 24 (2): 200–18.

Mayntz, Renate. 2004. "Mechanisms in the Analysis of Social Macro-Phenomena." *Philosophy of the Social Sciences* 34, no. 2 (June): 237–59.

Meegdenburg, Hilde van. 2019. "'We Don't Do That': A Constructivist Perspective on the Use and Non-Use of Private Military Contractors by Denmark." *Cooperation and Conflict* 54, no. 1 (March): 25–43.

———. 2023. "Process Tracing: An Analyticist Approach." In *Routledge Handbook of Foreign Policy Analysis Methods*, edited by Patrick A Mello and Falk Ostermann, 405–20. London: Routledge.

Runhardt, Rosa W. 2016. "Tracing the Productive Continuity of Social Mechanisms." Qualitative & Multi-Method Research 15, no. 1-2 (Spring/Fall): 22–8.

Qualitative and Multi-Method Research Spring 2023, Volume 21.1 https://doi.org/10.5281/zenodo.7921457

Ordinary Extraordinary Ethnography

Louisa Lombard
Yale University

This is a very useful book, for which ethnographers and other qualitative social scientists should be grateful. It is helpful in the way of style guides: a rejuvenating reminder of first principles, helpful no matter the stage of one's mastery of the subject. Qualitative researchers can easily fall victim to the "curse of knowledge" (Pinker 2014) and take shortcuts in both the analysis and presentation of data. Small and Calarco name five traits that sound qualitative research has: cognitive empathy, heterogeneity, follow-up, and self-awareness. Like a style guide that includes first a convoluted sentence and then a clear one, they include interview and participant observation extracts that do not benefit from the robustness that comes from those traits and then ones that do. The side-by-side comparisons are compelling and give practical ways for people to improve their literacy both in consuming and producing research.

At times, the authors' examples seem a bit too perfect. I have never interviewed anyone who spoke in the tidy, direct quotes the authors present as examples. I'm not sure anyone has. But reading them improved my research practice. Based on the authors' reminders about getting into the "hows" and the "whys" that are the added value of qualitative research, I found myself asking follow-up questions more attentively.