# QMMR
## Qualitative and Multi-Method Research
Fall 2023 | Volume 21.2


**Inside:**



ISSN: 2153-6767

QMMR Qualitative & Multi-Method Research

**Fall 2023 | Volume 21.2**

# Table of Contents

QMMR Qualitative & Multi-Method Research

**Fall 2023 | Volume 21.2**

# Table of Contents *(continued)*

## Book Symposium: Tasha Fairfield and Andrew Charman. 2023. *Social Inquiry and Bayesian Inference.* New York: Cambridge University Press



## Notes from the Field



## Notes from the Classroom

QMMR Qualitative & Multi-Method Research

# Letter from the Section President

Qualitative and Multi-Method Research Fall 2023, Volume 21.2 https://doi.org/10.5281/zenodo.10197156



I want to start this brief letter by extending a heartfelt thanks to all of you—our qualitative and multi-method membership—for your stalwart commitment to the rigorous application of qualitative and mixed methods in the discipline. I believe our work is needed now more than ever. The last few years have not been easy, and the future is daunting. We face drought, flooding, catastrophic storms, and other increasingly alarming signs that our planet is overextended and overworked. Countries are at war; the human cost has been appalling. The consequences of the global pandemic loom large. The list of challenges, both historical and contemporary, is long, and it seems to be growing.

Our work, I repeat, is necessary: To disentangle the multiple, complex causes that underpin these troubling phenomena. To give voice to historically and contemporarily marginalized communities. To develop grounded theories that speak to big questions. To revise, revisit, reframe. To bring empirically rich, nuanced, and even contradictory data to bear on our understanding of the world around us.

The QMMR section seeks to support its members in these scholarly tasks. One way to do this, we suggest, is to promote innovation in the use of epistemologically and methodologically diverse data. On this point, I am eager to announce that we have created the Qualitative Evidence Award, which rewards scholarship that takes an original approach to conceptualization or to the generation and use of qualitative data. Our section believes strongly in, as our bylaws now state, "exploring new questions in novels ways and older questions with fresh eyes." It is my hope that this award will appeal to younger scholars and amplify their work. The call will go out for the first time in 2025, and will be an annual award from that point on.

As many of you know, QMMR has long aspired to be a center of scholarly activity, where practitioners of qualitative and mixed methods can find colleagues, co-authors, support, and inspiration. While I believe we have made impressive gains in these goals over the years, it is nevertheless the case that QMMR has not been a methodological home for all. Over the past few years, the members of the Executive Council have worked diligently to make the section more diverse and more inclusive. On this point, I am incredibly grateful to our outgoing colleagues for all of their hard work in moving this objective forward—especially Alan Jacobs, as former President, and Veronica Herrera, as former Vice President. Both Alan and Veronica were steadfast in their commitment to making QMMR more diverse and inclusive. The section has greatly benefitted from their leadership. Alan, in particular, set the bar very high for me as incoming president.

Under their leadership the section introduced a Diversity, Equity, and Inclusion officer, whose primary mandate is to facilitate initiatives that promote and facilitate diversity within the section and also make underrepresented groups more visible in qualitative and multi-method research. Our current DEI officer is Roselyn Hsueh, a fantastic scholar at Temple University. Our updated bylaws also include a generalized commitment to promote diversity, equity, and inclusion in QMMR's membership, scholarship, and initiatives—a commitment in which I believe strongly and that will guide my work as section president.

A continued goal, for example, under my tenure will be to make the QMMR research community a more welcoming space for scholars that belong to groups that have been historically marginalized in our profession. Toward that end, we are introducing an additional award that draws attention to research that focuses on the historic and ongoing impacts of discrimination and exclusion, as well as the struggles for inclusion, in society. The Politics of Marginalization and Inclusion Award will recognize an outstanding qualitative or multi-method publication that explicitly engages with and contributes to our knowledge of these areas. As with the Qualitative Evidence Award, we will put out the call for this annual award following the next APSA Congress.

Additionally, the Emerging Methodologists Workshop, founded by Diana Kapiszewski and Hillel David Soifer in 2022 with support from the National Science Foundation, gathers annually six advanced graduate students and junior faculty from historically marginalized groups who are interested in developing, teaching, and publishing on techniques for qualitative or multi-method research. Workshop participants receive guidance and feedback from "Methods Mentors," helping them to produce a publishable-quality paper on collecting and/or analyzing qualitative data. It took me a long time to see myself as a qualitative methodologist, even though I had a lot of tools at my disposal—including a supportive dissertation committee and qualitative methods training. The EMW is designed to help younger scholars step into that role with greater confidence—to see themselves as valuable contributors to methodology regardless of their background or experience. The symposium in this issue provides a lot more information about this fantastic initiative, and it features the truly excellent work of the first cohort.

The New Voices Initiative at the Institute for Qualitative and Multi-Method Research (IQMR) at Syracuse University was designed to integrate outstanding early career researchers, including those from historically marginalized groups, into different teaching modules at the institute. The 2023 IQMR cohort was the first to benefit from these instructors, who brought new perspectives and insights on the method(s) they taught.

New and forthcoming publications offer opportunities for distinct voices to advance our understanding of methodology. Under the magnificent editorial leadership of Ezequiel and Juan, for example, this publication incorporated two new initiatives, which were first presented in the 2022 Fall issue (Volume 20.2). "Notes from the Field" promotes lessons and strategies acquired while in the field, while "Notes from the Classroom" promotes lessons and strategies from instructors on the ground. Both are open to younger scholars. In fact, the first two Notes from the Field featured the experiences and reflections of graduate students as they came back from the field. A forthcoming volume that Sara Wallace Goodman and I have edited[1] offers a comprehensive introduction to doing "good" qualitative research. All forty chapters are written by women and non-binary colleagues—individuals with years of experience doing qualitative work but whose voices have not been historically featured in our methodological texts.

To be sure, there is more work to do. But in these dark days, I find solace in our community and in its commitment to advance our understanding of the world, and to do so in a way that is thoughtful, open, and increasingly and intentionally inclusive.

Jenniffer Cyr
*Universidad Torcuato Di Tella*

1 *Doing Good Qualitative Research.* Forthcoming 2024. ISBN: 9780197633137

QMMR Qualitative & Multi-Method Research

# Letter from the Editors

Qualitative and Multi-Method Research Fall 2023, Volume 21.2 https://doi.org/10.5281/zenodo.10197826

In our third issue as QMMR editors we welcome Jennifer Cyr as the new president of APSA's Qualitative and Multi-Method Research section. As she points out in her inaugural Letter from the President, our community has taken important steps to promote diversity and inclusion. We would like to think of this edition of QMMR as an example of how the section's aspirations on this front can be put into practice.

In addition to our now regular Notes from the Field and Notes from the Classroom, the current issue features contributions by established and emerging methodologists with diverse backgrounds and experiences, including a taste of the material presented at the first edition of the Emerging Methodologists Workshop; a symposium on the legacies of an article that proved foundational for the interpretivist tradition; and another symposium evaluating a recent major contribution to the canon on causally oriented qualitative methods.

We hope this broad range of perspectives helps reinforce our community's mission statement and strongly encourages colleagues at all career stages and from all epistemological persuasions to join the section and submit their work to QMMR.

As Dvora Yanow notes in the introduction to the first symposium, Charles Taylor's magisterial "Interpretation and the Sciences of Man" turned fifty in 2021. This article played a central role in the development of interpretive methods in the social sciences. The essays featured here explain why. They variously discuss the importance of meaning-making for social and political explanation, critical questions regarding positionality, the centrality of diversity for the success of any research programme, and the perspective-shifting (or indeed paradigm-shifting) potential of interpretive work.

The second symposium celebrates a ground-breaking initiative spearheaded by Diana Kapiszewski and Hillel David Soifer, one that seeks to promote new voices in discipline-wide methodological debates. The Emerging Methodologists Workshop, which will meet once a year before the annual meeting of the American Political Science Association, invites early career researchers, including advanced doctoral students, to present papers that explore methods for qualitative data collection, analysis, and integration. The scholars invited to the first edition, whose work is featured in this issue of QMMR, delve into all aspects of the research process: questions related to fieldwork ethics, interview research, participant observation, mixed-method designs, measurement, and theory building. The essays leave no doubt that the future of the methods community is very bright.

Few methods books have been more eagerly awaited than Fairfield and Charman's *Social Inquiry and Bayesian Inference*. The book is already compulsory reading for all scholars who engage in qualitative research in general and process tracing in particular. The contributors to our third symposium rigorously evaluate the vision the authors set out for qualitative and mixed-method research. Fairfield and Charman then offer a comprehensive response to their critics.

The issue closes with our now regular *Notes from the Field* and *Notes from the Classroom*, which seek to draw practitioners and methods instructors into the QMMR community. Francesca Lessa's *Note from the Field* is a raw, yet analytically powerful discussion of what happens when the actors we study turn against the research project. She discusses her research on transitional justice in South America and how this led to her being the victim of anonymous threats while conducting fieldwork in Uruguay. Francesca walks us through the personal, ethical, and research challenges she had to confront in the aftermath of these threats, and what all of this means for engaged scholarship.

The *Notes from the Classroom* feature an instructor and one of his students unpacking an in-class exercise designed to teach research design. Specifically, Philip Ayoub and Jaya Duckworth outline the practicalities of an activity that invites students to reverse-engineeer a research design, discuss the pedagogical benefits of adding this type of formative assessment to our methods syllabi, and reflect on the obstacles all parties involved must overcome to make the exercise a success.

Before we let readers dig into this rich material, we would like to reiterate our usual call to all members of the QMMR community to submit original articles, symposia, and notes from the field and classroom for our consideration. Articles and symposia will be typically peer-reviewed, whereas we will review notes in-house. You can find details about submission guidelines on our website: https://www.qmmrpublication.com

Until the next issue,

Ezequiel Gonzalez-Ocantos
*University of Oxford*

Juan Masullo J.
*Leiden Unviersity*

QMMR Qualitative & Multi-Method Research

# Symposium: Revisiting the impact of Charles Taylor's "Interpretation and the Sciences of Man" after 50 years

Qualitative and Multi-Method Research Fall 2023, Volume 21.2 https://doi.org/10.5281/zenodo.8326478

## Revisting Charles Taylor's 1971 "Interpretation and the Sciences of Man"

Dvora Yanow
*Wageningen University*

> The hegemony of Cold War analytical philosophy of language was disrupted by…Charles Taylor's critique of atomist behaviorism through his magnificent essay on "Interpretation and the Sciences of Man"….
> —Seyla Benhabib (2020, 1047)

> A study that aims to analyze the dominant ethnicity from the bottom up must rely on qualitative methods that allow the researcher to examine the social reality from the subjective point of view of those who live within it (Taylor, 1987 [1971]).
> —Orna Sasson-Levy (2013, 34)

The year 2021 marked the 50th anniversary of the initial publication of Charles Taylor's influential article "Interpretation and the Sciences of Man." The article engaged how the human sciences might make sense of—interpret—meaning's multiple forms of expression, from language to acts to objects, and then evaluate that analysis. Taylor focused on "experiential" meaning—"for a subject," "of something," and existing only "in relation to the meanings of other things"—as distinct from linguistic meaning (1971, 11-12). The opening pages might be mistaken as a contribution solely to political theory or philosophy and the analysis of texts, except for the emphasis on the place of interpretation in the *human* sciences and on "text-analogues"—meaning-full acts (and, ultimately, objects) treated *as if* they were texts for analytic purposes—and their potential for manifold meanings.

Taylor's argument disrupted not only philosophy, as Benhabib (first epigraph) attests (pointing also to Rawls, Bernstein, and MacIntyre), but also methodological arguments concerning empirical research—what has since come to be called "interpretive"—and its legitimacy, as a phenomenological-hermeneutic undertaking. The second epigraph is an example of a common use of Taylor's article to defend a qualitative or interpretive research design. The purpose of this symposium is to celebrate the article's anniversary and reflect on what it has meant for interpretive political and social science. Contributors come from several subfields: public policy and administration, American government, comparative politics, and political theory. As background for considering its impact on methodological thinking and its standing in the discipline, this essay situates the article in the context of the social sciences at the time of its publication.

Appearing initially in the *Review of Metaphysics*, Taylor's article was quickly picked up in several important collections, becoming "one of the most widely-printed and widely-read articles ever published on the topic of interpretive social and political science" (David Forrest, personal communication, January 20, 2023): as of June, 2023, the original article and initial reprint had been cited over 3,600 times (according to Google Scholar; given its many reprintings, that number may well be higher). It is cited in many of the major works engaging interpretivism in political science (e.g., Bevir and Rhodes 2016, Lynch 2014, Schaffer 2016, Schatz 2009, Schram and Caterino 2006, Schwartz-Shea and Yanow 2012, Shenhav 2015, Yanow and Schwartz-Shea 2014). In recognition of Professor Taylor's contributions to advancing interpretive research in political and other social sciences, in particular through this article, the American Political Science

Association's Interpretive Methodologies and Methods Conference Group presents the Charles Taylor Book Award annually, for "the best book in political science that employs or develops interpretive methodologies and methods" (Interpretive Methodologies and Methods, n.d.).[1]

The article's language is, in some places, dated. The title, for instance, with "man" as a generic referent, reflects its times, and Taylor subsequently turned to the language of "human sciences" (e.g., Taylor 1985b). Otherwise, for those immersed in interpretive thinking, encountering some of the article's ideas in a 50th anniversary re-reading is like meeting old friends. Still, although it is anachronistic to expect a writer to use language and voice ideas that were not yet part of public or academic discourse when he wrote, such as "positionality" and "reflexivity," we might ask what, if any, considerations a reading 50 years later brings into high relief.

Symposium authors explore the article's meaning and relevance for interpretive-qualitative methodological thinking today. Among other things, they explore how we might understand its influence and whether its argument is still compelling for political science and its subfields, research methods, teaching, and/or advising students, and how Taylor's thought could be extended to meet present concerns: the need for political scientists "to think otherwise about their research practices," especially concerning the centrality to explanation of social actors' meaning-making (Peregrine Schwartz-Shea, this symposium); interpretive research's potential to engender a more "disruptive" political science, one which "highlight[s] possibilities for ordinary people to rework [the] world in their favor" (David Forrest, this symposium); the need for greater engagement with "embodied positionality," especially bearing on "women of color scholars and third world feminists" (Natasha Behl, this symposium); potential contributions of interpretive empirical research to political theory or political philosophy, especially concerning the relationship between expectations for "rule-guided," "systematic" human action and finding a "measurable," "predictable" world (Matt Longo, this symposium); "messy realit[ies] of political life" that challenge objective political science assumptions of order (Carolyn Holmes, this symposium).

Along with some of the other contributors to this symposium, I engage the article personally, remarking on when and how I first became aware of it, the significance of that context for my reading of it, and its impact on my thinking. Revisiting the article after so many years enables several additional sorts of reflections. For one, awareness of its existence can remind political scientists that *interpretive social science is not only not new*, but, indeed, *has important origins within the discipline*. It is not an interloper imported from elsewhere. Such a reminder can be significant for newer generations of scholars, in particular, especially in an era that has seen, and is still experiencing, arguments advancing various "innovations" from a kind of "unity of science" perspective—that is, as if all political and social sciences should be held to one, single standard reflecting one analytic method. That is precisely the argument that Taylor sought to counter in the article being celebrated here.

## Situating the Article

A text such as Taylor's might be situated in several ways. I briefly touch on where it stands in his own scholarly work and the article's publication history and then take up the methodological context in which it landed, as seen through other publications bracketing 1971.

## The Author

Born in Montreal, Charles Taylor earned two undergraduate degrees: in history at McGill, followed three years later by Politics, Philosophy, and Economics at Oxford's Balliol College. He took his DPhil at Oxford in 1961 working with Sir Isaiah Berlin. In the decade between receiving that degree and publishing the article in question, he brought out two books and several journal articles. Afterwards, he published, among articles and chapters, two books on Hegel; three relating social theory (or social philosophy) and the social sciences, a theme he returned to in his latest book; four on identity and multiculturalism; and one taking up religion (see references for details). Political activity—helping resettle hundreds of refugees in Vienna after the 1956 Hungarian revolt, delivering underground seminars in Czechoslovakia in the 1980s (Lukes 2018, 737)—and various awards—including as inaugural recipient of the prestigious Berggruen Prize in philosophy—were interspersed betwixt and between. The teaching and research areas listed at his McGill faculty page indicate the range of Taylor's interests and scholarship: Philosophy of Action, Philosophy of Social Science, Political Theory, Greek Political Thought, Moral Philosophy, the Culture of Western Modernity, Philosophy of Language, Theories of Meaning, Language and Politics, German Idealism (Abbey 2021; Calhoun 2016; McGill University, n.d.).

## The Article: Publication History

1 The award was proposed at the founding meeting of the Conference Group in Spring 2008; Ido Oren suggested naming it in honor of Professor Taylor's contributions.

Between 1973 and 2003 "Interpretation and the sciences of man" was reprinted seven times in collections edited by scholars in phenomenology, political science and philosophy, anthropology and philosophy, and philosophy of science or social science (Taylor 1971). It is the first chapter, for example, in both the 1979 and 1987 editions of Rabinow and Sullivan's influential *Interpretive Social Science* (which have a combined citation of nearly 2000). Taylor included it as Chapter 1 in his own *Philosophical Papers 2: Philosophy and the Human Sciences* (1985b, 15-57), which has garnered over 3400 citations.[2]

## A Broader Context of Ideas

Bracketing its initial 1971 appearance by a handful of notable publications brings the article's place in developing ideas into greater focus, explaining, too, perhaps, the demand to reprint it. Literary theorist Kenneth Burke's two major works—*A Grammar of Motives* (1945) and *A Rhetoric of Motives* (1950), republished in a single volume in 1950—drew out the performative implications of plays, in particular, for human action (see also Burke 1989), making the latter subject to analytic ideas developed in literary criticism, anticipating one of Taylor's key hermeneutic insights. Two publications in 1962 marked significant rifts in the landscape of thinking about knowledge and knowing, the one concerning language, the other, the practice of science. J. L. Austin's 1962 *How to Do Things with Words*, the 1955 William James lectures delivered at Harvard University, shifted the established understanding of language as descriptive alone to a sense of its performative dimensions, drawing a link between words (of a certain type) and what they enacted. "I do" (in a marriage ceremony) and "I christen thee" (spoken of a ship) accomplish more than just descriptions of events, Austin argued; they perform them (1962, 12). And Thomas Kuhn's *The Structure of Scientific Revolutions* showed that science does not proceed as had been thought, building brick by brick on previous, solo inventions. Instead, physical sciences are social practices, with new, "revolutionary" ideas developing when sufficient numbers of research anomalies can no longer be explained effectively through existing theories. These new ways of seeing would typically come from people on the margins of disciplines—women, for instance, or members of demographic minority groups—who had not yet been fully socialized to a discipline's dominant paradigm (1962, 209).

Attention to language was also beginning to grow in political science, often in reaction against the behavioralism that had captured it and other social sciences. Murray Edelman's *The Symbolic Uses of Politics* ([1964] 1985) focused on language and other symbols of government, implying a method of analysis of empirical materials drawing on the kind of hermeneutics that Taylor would develop a few years later. (In the 1985 edition's Afterword, Edelman was even more pointed in critiquing the positivist thinking that was his foil.) Two more path-breaking works followed, both in 1966. Chemist-cum-philosopher Michael Polanyi's *The Tacit Dimension* (1966) condensed his argument from the 1958 *Personal Knowledge* concerning science as a social practice, emphasizing its tacit dimensions: "We can know more than we can tell," he wrote (1966, 4; see also Nye 2011). And, bringing Schützian phenomenology to English-reading audiences, sociologists Peter Berger and Thomas Luckmann's *The Social Construction of Reality* (1966) articulated the processes through which intersubjective meaning-making becomes institutionalized.

The enlarged second edition of Kuhn's *Structure* was published in 1970, garnering a still wider readership. One year later, Taylor's article appeared in September as philosopher Paul Ricoeur's "The Model of the Text: Meaningful Action Considered as Text" was being published in *Social Research*. (With Taylor's, it, too, was picked up in the Dallmayr and McCarthy 1977 and Rabinow and Sullivan 1979 and 1987 collections.) The parallels between the two articles are strong, both of them arguing for the analytic utility of treating human action as if it were a text. Where Ricoeur's approach sought to join phenomenology and hermeneutics, Taylor's remained more grounded in language, treating acts as "text-analogues." That same year saw Edelman (1971) extend his explorations of the expressive dimensions of political acts in *Politics as Symbolic Action*. Two years on, anthropologist Clifford Geertz published *The Interpretation of Cultures* (1973), setting out an argument for seeing human action as expressive of meaning. Invoking empirical examples, he described the double hermeneutic of researchers developing their own interpretations of the interpretations made by situated actors in fieldwork settings. Capping off the decade in political science, Edelman's *Political Language* (1977) focused fully on that over the material objects and acts that had featured in his two previous books; and Richard F. Fenno's appendix "Notes on Method: Participant Observation" in *Home Style: House Members in Their Districts* (1978) detailed the kind of method Taylor's arguments justified.

This abbreviated publishing history points to the burgeoning interest across political and other social sciences in the ideas Taylor engaged in 1971: the characteristics and processes of a science of interpretation. A shift in textual interpretation may have been in the air, toward the phenomenology of "reader-response theory" (Iser 1989) and various dimensions of

2 I owe the push to feature citation counts to David Forrest.

the "linguistic turn." But in many respects this article lit the path, being all the more significant and laudable for advancing ideas of interpretation and meaning given the near-hegemonic command held by behavioralist thinking at the time. Its significance, as David Forrest notes, lies in its place as "a signal early publication from one of the most accomplished political and philosophical thinkers of the last half century, …[giving] voice to a much broader and more general transformation in academic thinking about interpretation…. [The article] reflects, exemplifies, [and] presages many of the most important insights to come out of that transformation" (personal communication, January 20, 2023), as the essays that follow make clear.

## One Meaning of the Text

Taylor's article got into my hands sometime between my 1982 dissertation and my first book (1996), reworking that dissertation. I know I had read it by 1988, when I had the good fortune to join the third-tier observers at an NEH summer seminar on interpretation. I learned a great deal from its "first-tier" who's who of hermeneutic, phenomenological, and practice studies scholars: seminar organizers Hubert Dreyfus and David Hoy and presenters Geertz, Kuhn, Taylor, Stanley Cavell, Alexander Nehamas, and Richard Rorty. (See Hiley, Bohman, and Shusterman 1991 for a set of seminar papers.) My memory is much sharper, however, on what the article meant to me at the time.

It constituted, first and foremost, validation of the line of argument that I was working to advance in the fields of public policy and organizational studies, concerning the centrality of meaning in human action and the place for expressing that meaning in more than instrumental-rational ways. The article's grounding in hermeneutics and the idea of "text-analogues" provided justification for going beyond literal legislative texts to include implementation acts treated as texts for purposes of analysis. Another of its important contributions was, in effect, a philosophical-methodological rationale for participant-observer ethnographic research in policy and organizational settings, including multi-sited ones.[3] In short: I had "fieldnotes"—diagrams as well as narrative "texts" and analytic thoughts, in some cases scratched on bits of paper; in others, in bound notebooks—recording observations, formal interviews, snippets of conversations, descriptions of events and interactions and the buildings and other spaces in which those had taken place, including who sat where and who spoke (or did not speak) to whom. These were my "translations" into words of what I had seen, heard, and experienced, useful reminders later, once I was out of "the field," of the details of those acts, events, interactions, and so forth as I tried to make sense of it all in relation to my research question. I had a method, in other words, or rather a handful of them, which I had used in situ, and I had started thinking about their methodological underpinnings. Taylor's article gave me ideas and language to think with, which were also useful in trying to explain to others how I was reasoning about what I had done.

That was important methodologically as both policy and organizational analyses had been colonized by instrumentally-rationalist accounts of human action, decentering previously engaged phenomenological-hermeneutic approaches. My political-organizational ethnographic work engaged the expressive, symbolic dimensions of human action and collective meaning. Taylor's critique of contemporaneous thinking in political and other social sciences encouraged me to consider not only that my inquiry was legitimate, but that I wasn't flying solo.

Looking back, I see three key contributions that the article made. First, as noted above, the idea of treating acts as "text-analogues" provided conceptual justification for extending analysis to acts in their own right and the material objects involved in these—including in policymaking and implementation, such as the built spaces in which acts and interactions take place (Yanow 1996). That idea lends itself to thinking not only about the fieldnotes that are significant in participant-observer and ethnographic research methods, but also about the transcripts that are central to interviewing research practices. Compare Taylor's "text-analogues" to Ricoeur's observation (1971, 530) concerning analysis: the object of exegesis includes not only writing, but also "the sorts of documents and monuments which entail a fixation similar to writing." The ideas are similar: Ricoeur's documents and monuments—which I inferred meant a form of built space—could be treated as we might literal texts, whether written by Senator Bernie Sanders or by an organization's Executive Director. Although the ideas run in parallel, Taylor's phrasing was

3 A distinction between "methodology" and "method" is key for understanding the discussion here and in some of the other essays in this symposium. Methodology is the "applied philosophy" of ontological and epistemological presuppositions that shape a research project and undergird particular research methods; methods put those presuppositions into practice. Some methods, such as ethnography, lend themselves to different methodological presuppositions; others, such as regression analysis, are inherently interpretivist or positivist in their presuppositions. See, e.g., Yanow ([2003] 2016), Yanow and Schwartz-Shea (2014, xxiii ff.). As David Forrest notes (personal communication, June 28, 2023), although ethnography might be used to generate "brute data" and assess mechanistic propositions about political behavior (although that would greatly diminish its promise), the methodological perspective that Taylor articulated permits an appreciation of its far more fruitful potential for exploring meaning-making.

more "operationalizable" for research practices.

Second, in his delineation of the meaning of "meaning," Taylor opened the door to a wider range of human traits than just instrumental rationality. Values, heartfelt beliefs, and feelings, not just "thinkings," have a central place in his philosophy and methodology. Third, his explication of how scientific analysis works insists on the intersubjective, societal dimensions of knowledge, much as Polanyi and Kuhn both argued concerning the practice of science and Geertz did with respect to the learning and transmission of culture.

Rereading the article now, I find other ideas for which at the time I lacked the "toeholds of the mind" (Sir Geoffrey Vickers, personal communication, January 1981) that would have enabled me to grasp them, but which I have since come to engage. One is the extent to which Taylor anticipated practice studies, especially phenomenological approaches to workplace and other practices such as Schatzki's, drawing on Heidegger and Dreyfus, which have influenced my own analyses (e.g., Yanow 2004, 2015; Yanow and Tsoukas 2009). Taylor subsequently developed those aspects of his thinking (1983, 1985b, 91-115). The second concerns the implications of parts of the article for writing. Taylor says, for example, "We cannot measure such [human, hermeneutic] sciences against the requirements of a science of verification…." This, it seems to me, points to the role of persuasion in scientific writing: that is, data alone do not compel a reader to accept an argument; instead, writing and content are intertwined. Studying both structure and components shows *how texts persuade*, including in their use of metaphor and story-telling, ideas that have also been developed in other fields (e.g., Brown 1976 and Gusfield 1976 in sociology; McCloskey 1985, 1994 in economics). Such approaches undergird the whole idea of writing as method (Schwartz-Shea and Yanow 2002, 2009) and the notion that in considering the reader of our texts, we are dealing with a third hermeneutic—that is, with readers' interpretations of our analytic interpretations of situational members' interpretations (Yanow 2009).

Taylor was not writing in a vacuum—meaning that his ideas were not the only ones to contribute to interpretive thinking, which was developing at the time, such that teasing his legacy apart from other thinkers' is nigh impossible. But three key factors (at least) have made his essay so significant: it came very early in the development of interpretive social science; the clarity of the argument and expression made it accessible; the stature of the author and his position among political theorists engaging empirical questions lent it added weight.

## In Closing

As I was contemplating the chronology surrounding Taylor's article, I realized that it was published the year I graduated college. I was not a good "Politics" undergrad: its concerns, as presented in my courses, did not resonate for me. How different my undergraduate "career" might have been, I thought, had the article appeared a bit earlier, especially in a political science journal, and circulated among the department's faculty (perhaps along with Murray Edelman's 1964 book). That might have meant, for me, a less circuitous route to graduate school, perhaps even a degree in political science itself—although I would have missed out on the experiences that grounded my understanding of those and other works when I did encounter them and which became the focus of my initial research and writings. Sadly, it remains the case that people in the discipline still challenge new scholars (and others) who choose to pursue interpretive empirical research. I have heard or read some comments—others have been reported to me—which communicate that such undertakings are not quite legitimate, not fitting the image of what empirical research in political science is "supposed" to look like.

The reflections in this symposium's essays suggest that Taylor's article still has relevance today, for its ideas as well as for the politics of science. Substantively, the article remains a forceful argument on behalf of a science of interpretation—one worth funding alongside other sciences by governmental agencies and foundations world-wide. It also voices a challenge to epistemological orientations that would banish interpretation from the human sciences (1971, 17). The article's critiques of political behavior studies and its engagement with the language of "brute data" may be a bit dated, but those ideas survive in the discipline under newer rubrics, such as in debates over replication. Other things have not changed so much. For example, "Intersubjective meanings…do not fit into the categorial grid of mainstream political science" (1971, 29)—still.[4] Finally, growing attention has focused of late on matters of truth, non-knowledge, and ignorance, in public and academic discourses alike. One implication of this development for methods "training" (and perhaps methodological argumentation) suggests "a shift away from traditional research strategies of reducing ignorance" towards those encompassing greater coping capacity when faced with ambiguities (Gross 2007, 745). The place of ineradicable ambiguity in human meaning-making was part of Taylor's argument in 1971 for recognizing the interpretive dimensions of political and social research. His article still gives those engaging such inquiry solid conceptual, theoretical, and methodological ground on which to stand.

4 On related points, see David Forrest's essay (this symposium).

## Acknowledgements

This essay revises comments introducing a roundtable on Taylor's 1971 article at the 2022 WPSA conference (March 10-12, Portland, OR), sponsored by the Interpretation and Methods Section. My thanks to David Forrest, Matt Longo, and Peri Schwartz-Shea for thoughts on previous drafts.

## References

Abbey, Ruth. 2021. "Charles Taylor." *Encyclopedia Britannica* https://www.britannica.com/biography/Charles-Taylor

Austin, J. L. 1962. *How to Do Things with Words*, edited by James O. Urmson. Cambridge: Harvard University Press.

Benhabib, Seyla. 2019. "For Jürgen Habermas on His 90th Birthday." *International Journal of Constitutional Law* 17 (4): 1047–49.

Berger, Peter L., and Thomas Luckmann. 1966. *The Social Construction of Reality*. New York: Anchor Books.

Bevir, Mark, and R.A.W. Rhodes, eds. 2016. *Routledge Handbook of Interpretive Political Science*. New York: Routledge.

Brown, Richard Harvey. 1976. "Social Theory as Metaphor." *Theory and Society* 3 (2): 169-97.

Burke, Kenneth. 1945. *A Grammar of Motives*. New York: Prentice-Hall.

Burke, Kenneth. 1950. *A Rhetoric of Motives*. New York: Prentice-Hall.

Burke, Kenneth. 1989. *On Symbols and Society.* Edited and introduced by Joseph R. Gusfield. Chicago: University of Chicago Press.

Calhoun, Craig. 2016. "Charles Taylor Has Reimagined Identity and Morality for a Secular Age." LSE blogs (October 19). https://blogs.lse.ac.uk/southasia/2016/10/19/charles-taylor-has-reimagined-identity-and-morality-for-a-secular-age/ (accessed October 28, 2022).

Edelman, Murray. 1971. *Politics as Symbolic Action*. Chicago: Markham.

Edelman, Murray. 1977. *Political Language*. New York: Academic Press.

Edelman, Murray. (1964) 1985. *The Symbolic Uses of Politics*, 2nd ed. Urbana: University of Illinois Press.

Fenno, Richard F., Jr. 1978. *Home Style: House Members in Their Districts*. Glenview: Scott, Foresman.

Geertz, Clifford. 1973. *The Interpretation of Cultures*. New York: Basic Books.

Gross, Matthias. 2007. "The Unknown in Process: Dynamic Connections of Ignorance, Non-Knowledge and Related Concepts." *Current Sociology* 55 (5): 742–59.

Gusfield, Joseph R. 1976. "The Literary Rhetoric of Science." *American Sociological Review* 41 (1): 16-34.

Hiley, David R., James F. Bohman, and Richard Shusterman, eds. 1991. *The Interpretive Turn*. Ithaca: Cornell University Press.

Interpretive Methodologies and Methods. n.d. "The Charles Taylor Book Award." Accessed November 4, 2022. https://connect.apsanet.org/interpretationandmethod/the-charles-taylor-book-award/.

Iser, Wolfgang. 1989. *Prospecting: From Reader Response to Literary Anthropology*. Baltimore: Johns Hopkins University Press.

Kuhn, Thomas S. (1962) 1970. *The Structure of Scientific Revolutions*, 2nd ed. Chicago: University of Chicago Press.

Lukes, Steven. 2018. "A Capacious Mind." *Philosophy and Social Criticism* 44 (7): 736–38.

Lynch, Cecelia. 2014. *Interpreting International Politics*. New York: Routledge.

McCloskey, Deirdre N. 1985. *The Rhetoric of Economics*. Madison: University of Wisconsin Press.

McCloskey, Deirdre N. 1994. "How to Do a Rhetorical Analysis of Economics, and Why." In *Economic Methodology*, edited by Roger Backhouse, 319-42. London: Routledge.

McGill University. n.d. "Charles Taylor." Accessed October 28, 2022. https://www.mcgill.ca/philosophy/people/emeritus-faculty/taylor.

Nye, Mary Jo. 2011. "Political Foundations of the Philosophies of Science of Popper, Kuhn, and Polanyi." In *Michael Polanyi and His Generation: Origins of the Social Construction of Science*, 223-57. Chicago: University of Chicago Press.

Polanyi, Michael. 1958. *Personal Knowledge: Towards a Post-Critical Philosophy*. Chicago: University of Chicago Press.

Polanyi, Michael. 1966. *The Tacit Dimension*. Chicago: University of Chicago Press.

Ricoeur, Paul. 1971. "The Model of the Text: Meaningful Action Considered as Text." *Social Research* 38 (3): 529-62.

Sasson-Levy, Orna. 2013. "A Different Kind of Whiteness: Marking and Unmarking of Social Boundaries in the Construction of Hegemonic Ethnicity." *Sociological Forum* 28 (1): 27-50.

Schaffer, Frederic Charles. 2016. *Elucidating Social Science Concepts*. New York: Routledge.

Schatz, Edward, ed. 2009. *Political Ethnography*. Chicago: Chicago University Press.

Schram, Sanford F., and Brian Caterino, eds. 2006. *Making Political Science Matter*. New York: NYU Press.

Schwartz-Shea, Peregrine, and Dvora Yanow. 2002. "'Reading' 'methods' 'texts.'" *Political Research Quarterly* 55 (2): 457-86.

Schwartz-Shea, Peregrine, and Dvora Yanow. 2009. "Reading and Writing as Method." In *Organizational Ethnography,* edited by Sierk Ybema, Dvora Yanow, Harry Wels, and Frans Kamsteeg, 56-82. London: Sage.

Schwartz-Shea, Peregrine, and Dvora Yanow. 2012. *Interpretive Research Design*. New York: Routledge.

Shenhav, Shaul R. 2015. *Analyzing Social Narratives*. New York: Routledge.

Taylor, Charles. 1964. *The Explanation of Behaviour*. London: Routledge & Kegan Paul.

———. 1970. *The Pattern of Politics*. Toronto: McClelland & Stewart.

———. 1971. "Interpretation and the Sciences of Man." *Review of Metaphysics* 25 (1): 3-51.

Reprinted in:

———. 1973 David Carr and Edward S. Casey, eds., *Explorations in Phenomenology*, 47-101. The Hague: Martinus Nijhoff.

———. 1977 Fred R. Dallmayr and Thomas A. McCarthy, eds., *Understanding and Social Inquiry*, 101-31. Notre Dame: University of Notre Dame Press.

———. 1979 Paul Rabinow and William M. Sullivan, eds., *Interpretive Social Science: A Reader*, 25-71. Berkeley: University of California Press.

———. 1985 Charles Taylor, *Philosophical Papers 2: Philosophy and the Human Sciences*, 15-57. New York: Cambridge University Press.

———. 1987 Paul Rabinow and William M. Sullivan, eds., *Interpretive Social Science: A Second Look*, 33-81. Berkeley: University of California Press.

———. 1994 Michael Martin and Lee C. McIntyre, eds., *Readings in the Philosophy of Social Science*, 181-211. Cambridge: MIT Press.1———. 1998 Elmer Daniel Klemke, Robert Hollinger, David Wÿss Rudge, and A. David Kline, eds., *Introductory Readings in the Philosophy of Science*, 110-27. Amherst: Prometheus Books.

2003 Gerard Delanty and Piet Strydom, eds., *Philosophies of Social Science: The Classic and Contemporary Readings*, 182-86. Philadelphia: Open University.

———. 1975. *Hegel*. Cambridge: Cambridge University Press.

———. 1979. *Hegel and Modern Society*. Cambridge: Cambridge University Press.

———. 1983. *Social Theory as Practice*. Delhi: Oxford University Press.

———. 1985a. *Philosophical Papers 1*: *Human Agency and Language*. Cambridge: Cambridge University Press.

———. 1985b. *Philosophical Papers 2: Philosophy and the Human Sciences*. Cambridge: Cambridge University Press.

———. 1989. *Sources of the Self: The Making of the Modern Identity*. Cambridge: Harvard University Press.

———. 1992. *Multiculturalism and the Politics of Recognition*. Princeton: Princeton University Press.

———. 1994. *Multiculturalism: Examining the Politics of Recognition*., edited by Amy Gutmann. Expanded ed. Princeton: Princeton University Press.

———. 1994. *Multiculturalisme: Différence et Démocratie*. Paris: Aubier.

———. 1995. *Philosophical Arguments*. Cambridge: Harvard University Press.

———. 2007. *A Secular Age*. Cambridge: Harvard University Press.

———. 2016. *The Language Animal*: *The Full Shape of the Human Linguistic Capacity*. Cambridge: Harvard University Press.

Yanow, Dvora. 1996. *How Does a Policy Mean? Interpreting Policy and Organizational Actions*. Washington: Georgetown University Press.

———. 2004. "Translating Local Knowledge at Organizational Peripheries." *British Journal of Management* 15 (S1): S9-S25.

———. 2009. "Dear Author, Dear Reader: The Third Hermeneutic in Writing and Reviewing Ethnography." In *Political Ethnography*, edited by Edward Schatz, 275-302. Chicago: University of Chicago Press.

———. 2015. "After Mastery: Insights from Practice Theorizing." In *The Emergence of Novelty in Organizations*, edited by Raghu Garud, Barbara Simpson, Ann Langley, and Haridimos Tsoukas, 272-317. Oxford: Oxford University Press.

———. (2003) 2016. "Interpretive Empirical Political Science: What Makes This Not a Subfield of Qualitative Methods." In *Qualitative Research in Political Science Volume 1*, edited by Joachim Blatter, Markus Haverland, and Merlijn van Hulst, 157-66. London: Sage.

Yanow, Dvora, and Peregrine Schwartz-Shea, eds. 2014. *Interpretation and Method*, 2nd ed. Armonk: ME Sharpe.

Yanow, Dvora, and Haridimos Tsoukas. 2009. "What is Reflection-in-Action? A Phenomenological Account." *Journal of Management Studies* 46 (8): 1339-64.

Qualitative and Multi-Method Research Fall 2023, Volume 21.2 https://doi.org/10.5281/zenodo.8326482

# From Philosophical Insight to Methodological Language: Charles Taylor, Interpretive Social Science, and Empirical Practice

Peregrine Schwartz-Shea
*University of Utah*

> Sometimes men show amazing prescience: the myth of Faust, for instance, which is treated several times at the beginning of the modern period. There is a kind of prophesy here, a premonition. But what characterizes these bursts of foresight is that they see through a glass darkly, *for they see in terms of the old language*: Faust sells his soul to the devil. They are in no sense hard predictions. Human science looks backward. It is inescapably historical.
>
> –Charles Taylor (1971, 50-51; emphasis added)

As might be expected given the place of political theory in the discipline of political science, Charles Taylor seems to be better known in that subfield than in the "empirical" subfields.[1] Yet in the article that is the topic of this symposium, he planted a seed that provided many of the insights that have since coalesced into a methodological language for imagining, designing, conducting, and writing up interpretive empirical political science projects—projects that, by definition, put the meaning-making of social actors at the center of explanation.

The epigraph provides inspiration for this way of reading the influence of Taylor's article. Its language is clearly of its time and for its intended audiences even as it introduced key interpretive ideas and anticipated others. After briefly considering the purpose of Taylor's piece in its disciplinary context, I consider the connections between it and the vocabulary available today to interpretive researchers. Some of these connections are readily evident, such as Taylor's use of the "hermeneutical circle,"[2] whereas others are implicit in his thinking, with subsequent developments producing new terms that summarize methodological thinking over the decades. There now exists, in a way that was not the case fifty years ago, an interrelated set of concepts,[3] a methodological language, that is appropriate to conducting interpretive social science and, also, to teaching what it means to do empirical, rather than philosophical, work (cf. Bevir and Blakely 2018).[4] Such methodological language has enabled the interpretive community to articulate to peer reviewers, editors, funding agencies, and others *how* its own scientific practices generate trustworthy knowledge.

## Taylor's Article as a Provocation

What might have been Taylor's primary purpose

1 The publication history in Yanow's symposium contribution shows this general pattern, with editors from philosophy or political theory outnumbering editors from the fields of anthropology, political science, and sociology. The division between political theory and the empirical subfields (i.e., international relations, comparative politics, American politics, among others) has been constructed historically (Kaufman-Osborn, 2006, considering U.S. political science) and bolstered by the fact-value dichotomy (Mihic, Engelmann, and Wingrove, 2005)—with the latter contested in interpretivism.

2 The "hermeneutic circle" examines how the meaning of the whole of a text cannot be grasped independent of its constituent parts (and vice versa). It characterizes the sense-making processes of interpretive research and, more broadly, of learning in general (Schwartz-Shea and Yanow 2012, 30-1).

3 For an overview, see Schwartz-Shea and Yanow (2012).

4 For an insightful take on the extent to which philosophers writing on social science do and do not read the relevant social scientific literature, see Becker (2011). I would add to Becker's point the importance of working with actual empirical evidence for understanding how philosophical presuppositions are enacted in methods practice (Schwartz-Shea 2019). Taylor himself did pay attention in his article to some of those methods practices; see his discussion of opinion surveys (19-22) in which he brings out their philosophical implications.

in writing his 1971 article? How should we understand what he was trying to accomplish? Of course, multiple inferences concerning his purpose are possible, as his article itself emphasizes. For me, a fruitful way to imagine that purpose is that it was intended as a *provocation* to mainstream empirical political scientists of his time.[5] It was asking them to think otherwise about their research practices, thereby challenging contemporary enthusiasm for behavioral political science. As a provocation (for a similar characterization see the essays of Forrest and Holmes in this symposium), the article did not do (and perhaps could not have done) a number of things that other readers (prominently, philosopher Martin 1994) seem to believe Taylor should have done: (a) address mainstream political scientists primarily in terms of what they saw as significant, such as their conception (at that time) of causality;[6] (b) speak primarily to the concerns of philosophers; (c) explicate for empirical political scientists the canon of relevant interpretive sources available at that time, such as Dilthey ([1900] 1990), Polyani (1966), Gadamer ([1960] 1976) and others. Such tasks were clearly beyond even an article as lengthy as Taylor's—close to fifty print pages.

Instead, what Taylor's article, seen as a provocation, did was to engage interlocutors' attention by making an argument that disrupted common narratives about what social science entails and about what it might mean to do empirical research differently from then available mainstream approaches. Understanding what is disruptive at a particular moment requires historical exegesis (some of which is available in Yanow's essay in this symposium). Mitchell (1991) provides an overview of the post-war political science context that demonstrates, in prominent scholars' own words, the "mission" and beliefs about what political science could accomplish. As one example, in 1944 Loewenstein wrote that comparative politics should become "'a conscious instrument of social engineering... [for] imparting our [U.S.] experience to other nations and…integrating scientifically their institutions into a universal pattern of government'" (quoted in Mitchell 1991, 79). While the blatancy of such sentiments may have abated by the 1970s, they are indicative of the challenges Taylor faced in proffering an alternative conceptualization of political science—a conception that directly challenges the possibility and desirability of any "universal [ahistorical] pattern" of governing.

## Taylor's Characterization of Mainstream Social Science

Taylor (1971) does not begin with a description of the mainstream approach to the human sciences. Rather, that portrayal is woven throughout the article, beginning in the second subsection of Part I in which he inquires about the stakes involved in advocating for a hermeneutical rather than an "empiricist" approach (7). Figuring prominently in that section is the concept of *brute* data—"data whose validity cannot be questioned by offering another interpretation or reading, data whose credibility cannot be founded or undermined by further reasoning" (8). Such data reflects the "highest ambition" of empiricism, that is, "to build our knowledge from such building blocks which can be anchored in a certainty beyond subjective intuition" (7). Put another way, data are "brute" to the extent that they are treated as foundational. The term thereby works to stop, intentionally or otherwise, the scholarly conversation, foreclosing the very hermeneutic circle about the topic under examination.[7]

Reading Taylor's definition now, I can imagine a reader objecting that scholars of all stripes *do* debate the merits of the data in any particular project. Yet Taylor's point about such *ambition* is palpable in contemporary claims about transparency, in which reader access to a scholar's evidence is envisioned as a means of arriving at agreement within and across epistemic communities or, at least, as a means toward more efficient "accumulation" or "transfer" of knowledge (Lupia and Elman 2014, 20). Admitting the hermeneutic circle into the human sciences challenges that ambition of finding evidence that can *settle* essential debates among scientists and within the polity (thereby obviating the politics of evidence and its

5 Among such mainstream empirical political scientists cited and discussed by Taylor, the most well-known are Gabriel Almond, Robert Dahl, and Seymour Lipset.

6 Martin takes Taylor's vagueness on issues that *Martin* believes are important as an opportunity to paint a portrait of Taylor's article, and the possibilities of interpretive social science, as overly narrow and constraining, a depiction I find ironic given the constraints of mainstream empirical political science of that time. Martin (1994, 26) takes Taylor's silence on causality as indicating that Taylor ~~(1994, 26)~~ "thinks causality is unimportant," but silence may mean a number of things, one of which Martin does not entertain—that Taylor chose not to address the mainstream solely on its terrain, ~~such as~~ that is, with its priorities and conceptualizations of science.

7 In this symposium, Forrest also discusses the meaning of "brute" or "brute data" at the time Taylor was writing, and Holmes gives several examples of how the pursuit of brute data remains a driving force for many researchers in the contemporary period. "Raw" now seems to have replaced "brute" in methodological discussions; see, for example, Pachirat's (2015) critique of the DA-RT (data access and research transparency) movement in political science.

interpretations).[8]

Second, more than halfway through the article (subsection iii, Part III), Taylor (1971, 27) introduces a phrase that is very odd to contemporary ears, "the categorial grid of behavioral political science."[9] His concern is that this particular scientific mode of thinking—a carving up of the experiential world into independent and dependent variables—forecloses understanding of the intersubjective phenomena that undergird political contestation and consensus.[10] In other words, his provocation to his interlocuters is that the behavioral grid misses a lot of what should be, in his view, part of the domain of a genuinely *social* science. (For extended examples of what gets missed, see Geertz's 1973 discussion of the distinctions between a blink, a wink, and a twitch; and Forrest, this symposium.) In the 1970s, for those investing in techniques that both depend on variables thinking and enable statistical treatments, Taylor's analysis called into question the emerging dominance of those approaches (Hauptmann 2022).[11]

Taylor's third characterization emerges further on in the article (subsection I, Part III) where he takes on the possibility of a "universal vocabulary of behavior" (1971, 33) that would enable the differentiation of societies without recourse to societal actors' meanings that emerge from their practices. The example of his day is functionalism—since discredited (Hawkesworth 2014)—but it is an impulse apparent, despite her own earlier field experiences, in the work of Nobel Prize winner Elinor Ostrom (1990): her research team traveled the world, translating local actors' conceptualizations and practices into universal rational choice categories, rendering indiscernible other ways of seeing and understanding "commons logic" (Schwartz-Shea 2010). Or as Taylor put it (1971, 33), the "categorial principles [of the verification model of political science make]…a whole level of study…invisible." That a universal scientific vocabulary would be an obstacle to knowledge is likely something Taylor's interlocuters had never considered, as Taylor expressly states: "the danger that such universality might not hold is *not even suspected* by mainstream political scientists" (34; emphasis added). Since Taylor's observation, interpretivist scholars have further challenged the universality of scientific vocabulary in innovative ways, contrasting "etic," "experience-distant," or "outsider" forms of knowing with "emic," "experience-near," or "insider" forms—introducing distinctive ways to think about theorizing and theory.[12]

In sum, Taylor's representation of the empirical political science of his time may well have provoked, even irritated, its adherents. Rather than engaging with them on their own terms—variables, causality, laws—he puts to them an essential question: what does your formulation of science miss? And instead of lauding universality, he pointed to its limitations for the human sciences.

## Explicit Interpretive Terminology

Taylor employs terminology, ideas, and arguments that elucidate philosophical perspectives which together point toward another way of doing empirical research, in contrast to the use of a categorial grid of universal concepts. The hermeneutic circle, of course, veers from the notion of a singular, objective "truth," instead seeing the human sciences in terms of a community of scholars. Subsequent thinking (Alston 1989; Knorr Cetina 1999) has further identified the diversity of such groupings—"epistemic communities"—that form pragmatically around shared topics, concerns, and/or presuppositions. This development has been essential to recognizing how persuasion operates, with particular ideas, arguments, and evidence being challenged, developed, and winnowed in ways that deepen knowledge—a hermeneutic "circle-spiral" (the phrase is Bentz and Shapiro's, 1998), if you will. Put differently, scholarly knowledge claims are not like the conspiracy theories of today's internet: epistemic communities assess, advise, and produce reasoned judgments through informal and formal peer review processes that discipline truth claims.

While the hermeneutic circle helps us to better understand scholarly practices (as opposed to appeals

8 The anxiety for some of admitting the hermeneutic circle into the social sciences became apparent, again, in the sometimes-feverish critiques of "postmodernism" that first emerged in the 1990s and continue to this day; see, for example, Read (2016).

9 Taylor never uses the word "quantitative," referring to variables-based research only obliquely through this phrase, "the categorial grid of behavioral political science"; in this way, his critique encompasses what is now termed positivist qualitative research. As David Forrest (this symposium) notes, researchers in that tradition (prominently, Brady and Collier, 2010, 318) aspire to qualitative data—termed "causal-process observations"—that are on equivalent footing to quantitative "data-set observations."

10 Whereas "consensus" is typically overtly recognized by actors within particular situations, intersubjective assumptions are usually unstated or tacit, setting the terms of the debate within which interlocuters engage.

11 Hauptmann's (2022) book-length treatment of the development of American political science traces the introduction and solidification of behavioralism in political science from the post-WWII period to 1970, providing telling historical context for Taylor's article. As Hauptmann documents, two of the scholars Taylor cites, Almond and Dahl, were key beneficiaries and proponents of what she calls a disciplinary "transformation" of U.S. political science.

12 For an overview and analysis of these term pairs, see Schwartz-Shea and Yanow (forthcoming).

to brute data to settle disputes), interpretive epistemic communities share, additionally, at least four ideas that are also evident in the article. First, Taylor's innovation[13] of the "text analogue"—as a human's accounting of conduct (as contrasted with measurement of variables)—opens up the forms of evidence available for scholarly analysis. Rather than being limited to quantitative data sets (based ineluctably on variables thinking), a panoply of possibilities can be entertained: policy documents, architectural spaces, discourses, field notes—none of which need be turned into numbers but which, instead, can be analyzed holistically. I can still remember my excitement years ago when I first understood empirical evidence in these terms. It felt liberating to understand I could legitimately generate data in a variety of forms (and that then led to me thinking about the form of *research questions* and how their particular formulations imply different kinds of evidence).

Second, Taylor's emphasis on intersubjectivity and its importance to the human sciences means that text and text-analogue evidence (in any of its forms) can be interrogated for traces of its human dimensions. What interactions and practices (by whom and when) produced this policy, discourse, or building? Rather than a research approach that focuses on adding up individual subjectivities (see Taylor's critique on page 27), intersubjectivity emphasizes that communities, societies, and civilizations require shared assumptions in order to do things that no individual can accomplish alone.[14] This was a major theme of Taylor's article, goading mainstream practitioners to consider what they might be missing from their research approaches since "[i]ntersubjective meanings…do not fit into the categorial grid of mainstream political science" (29). Any human science that makes no room for intersubjective processes of meaning-making renders invisible significant political phenomena.[15]

Third, intersubjectivity is borne of human collective practices enacted through language. Taylor implies that an understanding of language as mere labeling is insufficient; instead, he reiterates a constitutive understanding of language throughout the article, and one that is tied explicitly to practices. On this linkage, he states (1971, 24):

> There is no simple one-way dependence here. We can speak of mutual dependence if we like, but really what this points up is the artificiality of the distinction between social reality and the language of description of that social reality. The language is constitutive of the reality, is essential to its being the kind of reality it is. To separate the two and distinguish them as we quite rightly distinguish the heavens from our theories about them is forever to miss the point.

The import of this emphasis is, or should be, clear to us now: language is not epiphenomenal, something whose processes can be largely ignored by social scientists; rather, it is essential to grasping the potential incommensurability of different communities' ways of living.

The provocative nature of Taylor's argument—that the language-practice nexus may produce incommensurable understandings—is evident in one of Martin's (1994) critical assessments of Taylor's hermeneutic logic. Taylor's prose, which Martin examines, is worth quoting at length (47-48; emphasis added):

> The practical and the theoretical are inextricably joined here. It may not just be that to understand a certain explanation one has to sharpen one's intuitions, it may be that one has to change one's orientation—if not in adopting another orientation, at least in living one's own in a way which allows for greater comprehension of others. Thus, *in the sciences of man insofar as they are hermeneutical there can be a valid response to "I don't understand" which takes the form, not only "develop your intuitions," but more radically "change yourself." This puts an end to any aspiration to a value-free or "ideology-free" science of man.* A study of the science of man is inseparable from an examination of the options between which men must choose.

Quoting the italicized portion of this paragraph, Martin's characterization of Taylor's point here is that Taylor thinks that "interpretations are *ultimately* based on *rationally unsupported intuitions* and value decisions" (1994, 265; emphasis added). What Martin misses is that the practice of interpretive empirical social science (ethnography in particular) equips researchers to explore and understand the intuitions and value decisions of those whose world views differ from our own. The

13 Although Ricoeur introduced a similar idea in an article published in the same year, Taylor coined the memorable term, text-analogue. See Yanow's symposium essay for additional detail.

14 For elucidation of this point, see Berger and Luckmann's 1966 book, *The Social Construction of Reality*, where they analyze the emergence of collective, i.e., intersubjective, meanings. Thanks to Dvora Yanow for this point. Although Taylor does not cite the authors or use the book title's phrase, he embraces a constitutive logic of language that is consistent with their analysis.

15 For many policy examples of how interpretivist scholars surface community members' intersubjective assumptions, see Pader (2014). Without understanding such assumptions, analysis of conflicts *between* communities may be reduced to simplistic assessments of differing "interests."

constitutive logic of language is "world-making" (Nelson Goodman's 1978 term) in ways that escape the vocabulary and research logic of variables thinking and research. Research can change researchers—so that they are better able, contra Martin, to "rationally" explain others' ways of living for readers.

Fourth, Taylor describes human science as "inescapably historical" (51), as in the epigraph to my essay. This judgment informs his discussion of the hermeneutic circle, intersubjectivity, and language. It demands of scholars a recognition of how their own times and selves may limit what they do and learn, also reinforcing skepticism of the adequacy (and potential dangers) of any universal vocabulary or framework. This point has implications for "reflexivity," a contemporary methodological term of art to which I return below.

These ideas have been generative for the practice of empirical interpretive research. Just as positivist researchers approach a specified research question anticipating measurement of variables and considering causal relations and tests against data, interpretive researchers anticipate: learning in the field and thereafter (consistent with the hermeneutic circle); identifying texts and text-analogues relevant to their question; being attuned to language use (a constructivist ontology of language); and, ultimately, seeking to understand the intersubjectively constructed worlds of those they study. Taylor's ideas discussed thus far, then, fit together as part of an interpretive research *gestalt* (Schwartz-Shea and Yanow 2012, 113), but they leave some gaps that require additional unpacking of his article for a contemporary audience.

## Taylor's Implicit Ideas that have subsequently been developed

For those familiar with interpretive methodologies and methods, multiple places in Taylor's article anticipate ideas that have been more fully developed in the ensuing decades. As a first example, even though he names only traditional forms of scholarly logic (inductive and deductive), on his first page he lays out a conception of hermeneutics consistent with "abductive" logic (i.e., logic that makes sense of research puzzles). To wit, scholars confront "an object of study" that is "confused, incomplete, cloudy, seemingly contradictory"—a puzzle; hermeneutic interpretation "aims to bring to light an underlying coherence or sense" (1971, 3). This logic is reiterated in his discussion of "criteria for judgment within a hermeneutical science" (5):

> A successful interpretation is one which makes clear the meaning originally present in a confused, fragmentary, cloudy form. But how does one know that this interpretation is correct? Presumably because it makes sense of the original text: what is strange, mystifying, puzzling, contradictory is no longer so, is accounted for.

Whether or not the account will withstand interactions with other interlocuters, for researchers considering their evidence (whether in the field or at the desk), Taylor's description here evokes those "aha" moments when understanding "clicks" into place (for a formal examination of these processes, see Agar 1986). As Lisa Wedeen has observed,[16] all researchers can draw on any of these forms of logic—inductive, deductive, abductive. Abductive logic, however, seems to intertwine with hermeneutic sense-making in a particularly coherent way. Moreover, it is a logic well suited to empirical interpretive research, in which researchers seek to resolve puzzles through explanations that *take into account* the meaning-making context of those studied (Schwartz-Shea and Yanow 2012, 32).

Second, although the "practice turn" in social science was many decades away, Taylor ties practice to intersubjective meaning-making, anticipating the role of *tacit assumptions* as a focus of much interpretive research. Although "tacit" is not a word that appears in the article, in his example of the practice of banking, Taylor (1971, 24) emphasizes that observable human interactions ("negotiation," "bargaining in good or bad faith") depend on intersubjective agreements that often go unacknowledged, such as the entire conceptual edifice of what a bank is or does—an edifice constructed over time in the ways Berger and Luckmann (1966) theorize. For a contemporary example of the significance of understanding tacit assumptions, see Pader's (2014, 202) discussion of the clash of tacit world views in the interactions of Athabaskans (who expect the dominant person to take the conversational lead) with state social workers (who expect the Athabaskans to ask for help by clearly articulating their needs). By surfacing tacit knowledge, contemporary interpretive research makes visible, and explicable, many political phenomena in ways that Taylor may have, indeed, hoped would be the case.

Third, while Taylor speaks of "facts" and "values,"[17] he does not do so in terms of the fact-value dichotomy. As already quoted above, Taylor states (1971, 48): "This [practice-meaning linkage] puts an end to any aspiration to a value-free or 'ideology-free' science of man. A study of the science of man is inseparable from an

16 Wedeen made this observation in a panel presentation; although it stuck with me, I cannot recall the year or conference.

17 Much of this discussion involves developing his previous theme (that the grid misses intersubjectivity) by delving into the practices of opinion researchers, including asking about how questions have been designed and, ultimately, whom the knowledge is for.

examination of the options between which men must choose." Thus, Taylor anticipates the rejection of, what Nagel (1986) later called a "view from nowhere"—a critique subsequently developed by multiple feminist philosophers of science (e.g., Haraway 1988; Harding 1992; Longino 1990).[18] Taylor's article emphasizes a context-focused analysis of why humans act as they do, not seeking "general laws" (language he doesn't use) but explanations anchored in how people themselves live and how *they* understand their lives. Notably, the meaning and usefulness of the concept of generalizability is still being debated in interpretive work, indicating a methodological tradition of pushing boundaries within its own epistemic communities.

Astute readers will note the absence from the article of one key concept that has been prominent in interpretive methodological discussions over the last decades: "reflexivity." Methodological engagement with this concept developed rapidly with the acceptance of feminist philosophers' arguments that the identity of scholars matters for research—from the formulation of their questions to their generation of evidence in the field to their analyses and knowledge claims. (A few recent examples include Koinova 2017, and Soedirgo and Glas 2020). It is now *de rigueur* for interpretive scholars to actively reflect on and acknowledge how their multifold identities—including assumptions based on their theoretical and community commitments—impact their research. Nothing this explicit seems to exist in Taylor's 1971 article; yet it *is* there in the beginning section as a "condition [interpretive social science] must meet", i.e., acknowledging "a subject for whom these meanings are" (4). As he expresses it later (10), meaning is always "for a person."

It is this tenet—meaning is always for a person, *and scholars, too, are people*—that complements Taylor's emphasis on the constitutive nature of language. Together, these two ideas clarify how and why interpretivism is distinctive, how its contributions are enabled by these philosophical presuppositions, and how "integrating" or "synthesizing" them with variables-based thinking is problematic. To make this point with students, I give them two diagnostic questions to assess whether an empirical research project *is* interpretivist (in contrast to the often, misleading indicator of whether the evidence is quantitative or qualitative):

1. What is assumed about the role of language? (Is it ontologically constructivist or realist?)

2. What is assumed about the role of the researcher? (Does the researcher aim toward objectivity or reflexivity?)

Embracing *both* answers to these questions is philosophically incoherent, clarifying the profound implications for empirical research of interpretivist presuppositions. Contemporary scholars can now better appreciate, and articulate to others, what Taylor was working to get his interlocuters to recognize—that research can be done *otherwise*, as something distinct from the variables-based "grid" that then so dominated the discipline.

As interpretive empirical research, and its methodological entailments, have developed, ideas that were implicit in Taylor's thinking—abductive logic, tacit meanings—have become part of an explicit methodological vocabulary. Taylor's rejection of the fact-value dichotomy (although he didn't call it that) is evident in his insistence that meaning-making cannot be considered in isolation from the identity of specific meaning-makers. It is always "for a person"—a scholar who perforce has particular values. These ideas are now expressed in somewhat different language, taking into account other developments and literatures. Specifically, the concerns about bias that surround a commitment to objectivity have been replaced for interpretivists with efforts to actively reflect on how embodied, positioned researchers generate and analyze data and make knowledge claims. The meaning-making of those studied is now accompanied by analysis of the sense making of the scholars themselves. Scholarly interpretations of social actors' interpretations are then vetted in ongoing hermeneutical debates on substantive topics.

## Conclusion

Some fifty years later, my reading of Taylor's article shows that it succeeded as an invitation to do empirical political science otherwise. And that invitation has been taken up within the discipline such that interpretivism and interpretive empirical projects are now recognized in departmental curricula, journals (see, e.g., APSR, 2022), and awards and prizes. Even as some of his discussion was cast in the language of the 1960s and 1970s, Taylor showed amazing prescience. His provocations were sufficiently attention-getting that they were reprinted again and again, inspiring various scholarly readers and planting seeds that guided them in new directions. Thanks to Taylor and the developments he inspired, contemporary interpretive researchers can confidently

18 For me, personally, the fact-value dichotomy collapsed as I read the feminist philosophers cited here; my skepticism of it also emerged in methods teaching as students and I examined how different definitions of rape not only affected tabulations of incidence rates but also had clear political ("value") implications about who or what gets excluded and included in definitions and measures. As Behl's essay in this symposium suggests, Taylor's underdevelopment of ideas concerning scholarly identity may have made his essay less relevant to the thinking of younger generations of interpretivists.

claim interpretive social *science* and take their place in a more methodologically pluralist discipline.


## Acknowledgements

A rough version of these ideas was presented at a roundtable on Taylor's 1971 article held at the 2022 WPSA conference in Portland, Oregon, sponsored by the Interpretation and Methods Section. Thanks to Dvora Yanow for incisive feedback on a more fully-fleshed-out draft and to Matt Longo and an anonymous reviewer for useful comments on that draft.


## References


Agar, Michael H. 1986. *Speaking of Ethnography*. Newbury Park: Sage.

Alston, William P. 1989. *Epistemic Justification: Essays in the Theory of Knowledge*. Ithaca: Cornell University Press.

APSR Editors. 2022. "Publishing Your Qualitative Manuscript in the APSR." Available at https://www.cambridge.org/core/blog/2022/03/03/publishing-your-qualitative-manuscript-in-the-apsr/

Becker, Howard S. 2011. "Review of *Social World: The Structure of Human Civilization* by John R. Searle and *Fear of Knowledge: Against Relativism and Constructivism* by Paul A. Boghossian." *Science, Technology, & Human Values* 36, no. 2 (March): 273-9.

Bentz, Valerie M., and Jeremy J. Shapiro. 1998. *Mindful Inquiry in Social Research*. Thousand Oaks: Sage.

Berger, Peter L., and Thomas Luckmann. 1966. *The Social Construction of Reality*. New York: Anchor Books.

Bevir, Mark, and Jason Blakely. 2018. *Interpretive Social Science: An Anti-Naturalist Approach*. Oxford: Oxford University Press.

Brady, Henry E. and David Collier, eds. 2010. *Rethinking Social Inquiry: Diverse Tools, Shared Standards*. Lanham: Rowman & Littlefield, 2nd edition.

Dilthey, Wilhelm. (1900) 1996. *Selected Works*. Edited by R.A. Makkreel and F. Rodi. Princeton: Princeton University Press.

Gadamer, Hans-Georg. (1960) 1976. *Philosophical Hermeneutics*. Translated and edited by David E. Linge. Berkeley: University of California Press.

Geertz, Clifford. 1973. *The Interpretation of Cultures*. New York: Basic Books.

Goodman, Nelson. 1978. *Ways of World-Making*. Indianapolis: Hackett.

Haraway, Donna. 1988. "Situated Knowledges." *Feminist Studies* 14 (3): 575-99.

Harding, Sandra. 1992. "Rethinking Standpoint Epistemology." *Centennial Review* 36: 437-70.

Hauptmann, Emily. 2022. *Foundations and American Political Science: The Transformation of a Discipline, 1945-1970*. Lawrence: University Press of Kansas.

Hawkesworth, Mary. 2014. "Contending Conceptions of Science and Politics: Methodology and the Constitution of the Political." In *Interpretation and Method: Empirical Research Methods and the Interpretive Turn*, edited by Dvora Yanow and Peregrine Schwartz-Shea, 27-49. Armonk: M.E. Sharpe.

Kaufman-Osborn, Timothy. 2006. "Dividing the Domain of Political Science: On the Fetishism of Subfields." *Polity* 38, no. 1 (January): 41-71.

Knorr Cetina, Karin. 1999. *Epistemic Cultures: How the Sciences Make Knowledge*. Cambridge: Harvard University Press.

Koinova, Maria. 2017. "Beyond Statist Paradigms." *International Studies Review* 19, no. 4 (December): 597-621.

Longino, Helen. 1990. *Science as Social Knowledge*. Princeton: Princeton University Press.

Lupia, Arthur, and Colin Elman. 2014. "Openness in Political Science: Data Access and Research Transparency." *PS: Political Science & Politics* 47, no. 1 (January):19-42.

Martin, Michael. 1994. "Taylor on Interpretation and the Sciences of Man." In *Readings in the Philosophy of Social Science*, edited by Michael Martin and Lee C. McIntyre, 259-79. Cambridge: MIT Press.

Mihic, Sophia, Stephen G. Engelmann, and Elizabeth Rose Wingrove. 2005. "Making Sense in and of Political Science: Facts, Values, and 'Real' Numbers." In *The Politics of Method in the Human Sciences: Positivism and Its Epistemological Others,* edited by George Steinmetz, 470-95. Durham: Duke University Press.

Mitchell, Timothy. 1991. "The Limits of the State: Beyond Statist Approaches and Their Critics." *American Political Science Review* 85, no. 1 (March): 77-96.

Nagel, Thomas. 1986. *The View from Nowhere*. New York: Oxford University Press.

Ostrom, Elinor. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge: Cambridge University Press.

Pachirat, Timothy. 2015. "The Tyranny of Light." *Qualitative & Multi-Method Research* 13, no. 1 (Spring): 27-31.

Pader, Ellen. 2014. "Seeing with an Ethnographic Sensibility: Explorations Beneath the Surface of Public Policies." In *Interpretation and Method: Empirical Research Methods and the Interpretive Turn*, edited by Dvora Yanow and Peregrine Schwartz-Shea, 194-208. Armonk: M.E. Sharpe.

Polanyi, Michael. 1966. *The Tacit Dimension*. Chicago: University of Chicago Press.

Read, Rupert. 2016. "Richard Rorty and How Postmodernism Helped Elect Trump." *The Philosophers' Magazine,* November 22, 2016. https://www.philosophersmag.com/opinion/147-richard-rorty-and-how-postmodernism-helped-elect-trump.

Schwartz-Shea, Peregrine. 2010. "Elinor Ostrom's Fieldwork Sensibility," in "Beyond the Tragedy of the Commons: A Discussion of *Governing the Commons: The Evolution of Institutions for Collective Action*," symposium, *Perspectives on Politics* 8, no. 2 (June): 587-90.

Schwartz-Shea, Peregrine. 2019. "Scholarly Reflexivity, Methodological Practice, and Bevir and Blakely's Anti-Naturalism." *Critical Review: A Journal of Politics and Society* 31 (3-4): 462-80.

Schwartz-Shea, Peregrine, and Dvora Yanow. 2012. *Interpretive Research Design: Concepts and Processes*. New York: Routledge.

Schwartz-Shea, Peregrine, and Dvora Yanow. Forthcoming. "'Insiders' and 'Outsiders' in Field Research: An Interpretive Epistemological Understanding." In *The Oxford Handbook of Engaged Methodological Pluralism,* edited by Janet Box-Steffensmeier, Dino P. Christenson and Valeria Sinclair-Chapman. Oxford: Oxford University Press.

Soedirgo, Jessica, and Aarie Glas. 2020. "Toward Active Reflexivity." *PS: Political Science & Politics* 53, no. 3 (July): 527-31.

Taylor, Charles. 1971. "Interpretation and the Sciences of Man." *Review of Metaphysics* 25, no. 1 (September): 3-51.

Qualitative and Multi-Method Research | Fall 2023, Volume 21.2 | https://doi.org/10.5281/zenodo.8326493

# The Contributions of Charles Taylor's "Interpretation" Article: What They Are, How We Can Build on Them

David Forrest
*Oberlin College*

Between the late 1960s and early 1970s, several North American political scientists published essays aimed at what had become a glaring problem with the mainstream of their field —namely, its utter failure to recognize and make sense of the political and social turmoil of the previous decade (see, e.g., Wolin 1969; McCoy and Playford 1967; Taylor 1971). On the one hand, as these authors correctly argued, many of their colleagues depicted post-World War II politics and society—particularly in the United States—as fundamentally stable, well-functioning, and fair. On the other hand, this "world show[ed] increasing signs of coming apart" (Wolin 1969, 1081). In the United States, influential right-wing political formations such as the John Birch Society had challenged the basic tenets of postwar liberalism; in different ways, so had many ostensibly left-wing formations, such as the disruptive movements and riots that had swept across several major cities. How, these authors asked, could mainstream American political science have so miserably failed to engage with these happenings? And how could future scholars do better?

Given the sheer importance of their questions, it is perhaps unsurprising that a number of these authors' publications left major imprints on the discipline. The most influential among them showed how political science's *particular* failure vis-à-vis the 1960s revealed more *general* problems with its dominant approaches to research. Consequently, they inspired many future scholars to explore alternatives. For example, Sheldon Wolin's 1969 article "Political Theory as a Vocation" linked the discipline's failure to its behavioralist obsession with methodological "technique" and, as a result, became a touchstone for those invested in more critical theoretic scholarship. Likewise, the chapters published in the 1967 volume *Apolitical Politics*, edited by Charles McCoy and John Playford, traced the failure to the political biases of mainstream behavioral research and, in turn, helped to inspire a tradition of more avowedly progressive scholarship (Barrow 2008, 218-22).

But for those of us invested in the qualitative and, in particular, interpretive traditions of political science, the most significant contribution to this discussion was Charles Taylor's "Interpretation and the Sciences of

Man," published in 1971. Taylor's article took issue with many of the same intellectual shortcomings highlighted by scholars like Wolin and the contributors to *Apolitical Politics*. And like those other scholars, he argued that these shortcomings necessitated not just a revised view of the 1960s but a more general shift in the study of politics. However, he went much further in explaining and showing how the field's dominant approaches to research specifically foreclosed crucial examinations of *meaning-making*—the central focus of interpretive scholarship. In particular, as I argue in the next section, he showed how this approach stifled inquiries into the sociopolitical contingencies and tensions engendered by meaning-making—contingencies and tensions often overlooked but felt deeply in moments like the turmoil of the 1960s.

For two reasons, the current moment strikes me as a good one for taking stock of Taylor's article and what it contributed to political science's qualitative research tradition. First, we are now at a point where Taylor's insights have had over fifty years to percolate the field. They have influenced multiple generations of scholars, including many who have never actually read his article.[1] As a result, we can more confidently begin to assess the kinds of scholarship that his insights can enable and, in many cases, have enabled. We can better appreciate the proverbial fruits of his intellectual labor.

Second, we are currently amid another moment in which "the world shows increasing signs of coming apart" (Wolin 1969, 1081). Dominant social and political institutions in myriad countries have recently faced intense challenges from right- and left-wing movements. In the United States, for example, we can point to Donald Trump's and Bernie Sanders's presidential campaigns, the summer 2020 uprising that occurred in response to the police killing of George Floyd, the January 6, 2021, attacks on the US Capitol, a resurgence in labor protests and strike activity, and many others. As with the 1960s, this current moment pushes us to assess how well our approaches to research can make sense of politics and society, in all their messiness. It throws the general promise and limits of these approaches into starker relief.

Against this backdrop, the rest of this symposium article offers some brief reflections on two questions. The first is how Taylor's "Interpretation" article has left the field of political science better positioned to analyze politics and society. In short, what was its contribution? The second is about the limitations of his article. How might we extend his insights in ways that he either did not or maybe could not, given the historical and professional context in which he was writing?[2]

## Contributions

To my mind, the major contribution of Taylor's article was not simply to demonstrate that mainstream, North American political science inhibited interpretive examinations of meaning-making. It also explained and showed how this inhibition was wrapped up in a broader failure to analyze important contingencies and tensions in the arrangement of politics and society. More specifically, he showed how common but problematic notions about the ultimate *goals* of political science research and the proper *means* for achieving those goals pushed sociopolitical contingency and its relationship to meaning-making off the scholarly agenda.

Regarding goals specifically, Taylor heavily critiqued the field's taken for granted emphasis on what I would call precision—by which I mean the reduction of uncertainty about politics and society. Mainstream behavioral and, in Taylor's parlance, "empiricist" researchers in the 1960s focused intently on this goal, seeking not just factually accurate and theoretically compelling depictions of the world but depictions that "achieved certainty beyond subjective intuition" and made the world more predictable for readers (Taylor 1971, 7).[3] Even today, much of the discipline still treats the quest for precision or reduced uncertainty as the main—often the only—goal of political science research (Forrest 2016; Forrest 2017). As one prominent political scientist asserted in 2014, the discipline's highest aspiration should be to develop "technically precise analyses of the past" that "significantly clarify the future implications of current actions" and allow for "ever-increasing effectiveness and efficiency" (Lupia 2014, 6).

This single-minded focus on precision, Taylor suggested, is largely what leads political scientists to overlook important points of sociopolitical tension and contingency, particularly those that arise out of the ongoing and often unconscious fashioning of meaning. By pushing so hard to reduce uncertainty, precision-oriented research underplays these points, which are important because they render the world *more* uncertain and *less* predictable.[4] In his words, "this whole level of study"—that is, "the study of our civilization in terms of its intersubjective and common meanings" and the

1 Indeed, before preparing these reflections, I had never read his article (as far as I can remember), despite having published a book and multiple articles that engage with interpretive methods. For evidence of Taylor's overall influence, see Dvora Yanow's contribution to this symposium.

2 Regarding this second question, see also Carolyn Holmes's and Natasha Behl's contributions to this symposium.

3 See also Wolin (1969).

4 On this point, see also Matthew Longo's discussion of Taylor and sociopolitical "complexity" in this symposium.

contingencies those meanings entail—"is made invisible" (1971, 33).

To demonstrate, he gives the example of negotiations. In dominant North American institutions, he argues, negotiations follow a highly established and "very contractual" format (23). This format assumes the existence of distinct and autonomous parties and uses language like "entering into negotiation, breaking off negotiations, offering to negotiate, negotiating in good (bad) faith, concluding negotiations, making a new offer, etc." (22). Because this format is so established, in situations where it reigns, precision-oriented researchers can develop fairly certain accounts of how negotiations work—when they emerge or breakdown, why some parties succeed more than others, and so on. But in doing so, Taylor shows, they also distract from contingencies and tensions related to the format—the meaning of negotiation—itself. As he says, "other societies have no such conception" or format (23). So, how did it come about? Why is it so taken for granted? How might it decompose? Why is it so much less prominent in many other societies? Taylor demonstrates that any social science field taking precision as its primary or sole goal will fail to meaningfully address these kinds of questions.[5] Especially in a world that is "coming apart" and, thus, particularly rife with contingency and uncertainty, this cannot do.[6]

Regarding the means political scientists use in their quest for precision, Taylor focused his critical attention on the acquisition and analysis of what he called "brute data." By "brute data," Taylor meant "data whose validity cannot be questioned by offering another interpretation or reading" (8). In other words, "brute data" are empirical observations with relatively uncontested and standardized meanings (e.g., in a congressional roll call vote, a "yes" means support for bill passage, and a "no" means opposition). Around the time of Taylor's writing, many prominent political scientists championed the acquisition of this type of data, as it provided the only means by which to verify their precise claims about politics and society (see, e.g., Polsby 1960).[7] And still today, even among qualitative researchers, much of the field prioritizes the analysis of such "brute" observations.[8]

More effectively than anyone else at the time, Taylor articulated how this tendency—alongside the emphasis on precision—stifled analysis of the sociopolitical contingencies and tensions entailed by meaning-making. To put his main point simply: If all you do is acquire "brute data," then you can only study aspects of politics and society that are "brute data identifiable" (1971, 29). Sociopolitical phenomena whose meanings are especially contested—and, thus, particularly prone to becoming the source of contingency and tension—become unobservable. Taylor gives the example of legitimacy, or the extent to which "some societies enjoy an easier, more spontaneous cohesion which relies less on the use of force than others" (35). It is, Taylor demonstrates, impossible to observe legitimacy through the lens of "brute data." Here, he is worth quoting at length:

> "Legitimacy" is a term in which we discuss the authority of the state or polity, its right to our allegiance. However we conceive of this legitimacy, it can only be attributed to a polity in the light of a number of surrounding conceptions – e.g., that it provides men freedom, that it emanates from their will, that it secures them order, the rule of law, or that it is founded on tradition, or commands obedience by its superior qualities. These conceptions are all such that they rely on definitions of what is significant for men in general or in some particular society or circumstances, definitions of paradigmatic meaning which cannot be identifiable as brute data. (35)

Consequently, to observe and explain something like legitimacy, one *must* wade into the realm of hotly contested meanings and study the processes by which those meanings are made, reproduced, upended, etc. There is no other choice. The only way to remain in the universe of "brute data" is to instead treat legitimacy as a subjective state, an individual's opinion about a polity (perhaps captured in an interview or survey response). As Taylor puts it, "What enters into scientific consideration is thus not the legitimacy of a polity but the opinions or feelings of its member individuals concerning its legitimacy" (36). Again, especially in a moment such as the current one—where phenomena like legitimacy are

5 One could easily apply Taylor's critique to precision-oriented studies of other important sociopolitical phenomena. See, for example, James Ferguson's (1994) critique of "development" studies.

6 Something Taylor does not point out, but I have argued elsewhere, is that even some interpretive research adheres to the kind of precision-oriented, "empiricist" agenda that he associates mostly with "behavioral" or "mainstream" political science. This kind of interpretive research emphasizes its ability "to reduce uncertainty about how shared practices and understandings [or meanings] constitute society" (Forrest 2016, fn. 4). In other words, it focuses on meaning-making and its consequences—and, in that sense, moves beyond the strictures of "behavioral" political science—but still underplays the points of contingency and tension that meaning-making entails.

7 Even many of the behavioralists' critics also prioritized the acquisition of "brute data." For example, as John Gaventa explained in his foundational book *Power and Powerlessness* (1980), many scholars who criticized behavioralists' claims about political power nevertheless retained the behavioralists' insistence on observing power via observations with uncontested meanings.

8 See, for example, Brady and Collier (2004) and George and Bennett (2004), both of which are (deservedly!) still widely read and cited.

in flux and even harder to properly observe—such an approach cannot stand.

Taylor's critique of the quest for precision and the acquisition of "brute data" did not, of course, force either to the margins of political science. Nor did it prove that precision and "brute data" are somehow a total waste of time; many projects operating on the terms he critiques have produced and continue to produce valuable knowledge. What he did do, however, is underscore the need for scholarship—especially interpretive scholarship—that pushes beyond those terms. The disciplinary spaces where his insights gained more traction are also those where conceptions of political science have most fruitfully expanded, making the field more comfortable for research projects whose theories are not so precise or whose data is not so "brute."

One such space is the informal community of scholars who use ethnographic immersion to study politics (Schatz 2009). My recent book, which addresses the failures and egalitarian potential of social justice organizations in the United States, is an example of a project that has come out of this space (Forrest 2022). Far from striving for precision, the book intentionally aims to underscore points of contingency, tension, and unpredictability in the work of social justice organizing – points where organizers might create and seize opportunities to reach their greatest potential as democratic representatives. In addition, many of the phenomena I study—through observations of strategy meetings, protests, door-knocking, etc.—are avowedly *not* "brute." One could persuasively characterize these phenomena (one of which is actually legitimacy) in myriad ways. And to make sense of them and their explanatory importance, I had to examine how each one was contested, constructed, and made meaningful for different groups of people. Without Taylor's insights, one could easily treat these dimensions of my research as shortcomings. With his insights, however, it becomes clear that, in practice, they are enabling. They allowed me to examine a central process—what I call the fashioning of "contentious identities"—that (a) has major implications for whether social justice organizations develop a powerful and egalitarian voice on behalf of their constituents and (b) is largely invisible to more precise and "brute" scholarship.

## Extensions

By my reading, the contributions of Taylor's "Interpretation" article are considerable, and part of me wants to just dwell on and appreciate those contributions. Nevertheless, as qualitative political scientists, we should also reflect on how we might extend them. One way we can do this is by more explicitly articulating what political science research should look like once we have dethroned the quest for precision and the acquisition of "brute data." What else should qualitative scholars strive for, if not precision (in the sense of reduced certainty)? And how can we pursue that alternative goal, if not by gathering "brute data"? Taylor's article offers some initial and insightful answers to these questions. But it also leaves much unsaid, much for future scholars to elaborate.

When it comes to goals, the main alternative to precision that Taylor's article identifies can be summed up with the word provocation.[9] More specifically, as he says in the last section of his article (1971, 47-8), political science should provoke readers to "develop your intuitions" and, if necessary, "change yourself" "in a way which allows for greater comprehension of others," that is, other "fundamental options in life" or, I would say, sociopolitical possibilities. And it can do this, he suggests, by locating and investigating those same meaning-laden points of contingency—such as the "contractual" format underlying many North American negotiations—that precision-oriented research tends to underplay. Examining these points challenges readers to better appreciate how different "terms," or patterns of meaning-making, have, can, and might "structure the world in ways which are utterly different from, incompatible with our own" (47). In other words, rather than chase the "radically impossible" dream of "exact prediction" and certainty, such an examination pushes readers to develop an expanded sense of social and political possibility, rendering the world more unpredictable and uncertain (48-9).

For me, Taylor's goal of provocation is on the right track. However, as Taylor articulates it, it feels too vague, too open-ended. Provocation per se can push audiences in any number of directions, to unearth any number of possibilities, not all of which are equally worth our attention. Upon reading Taylor's article, we would be right to wonder, "develop your intuitions" and "change yourself" how? To what end?

As I have argued elsewhere, I think that a more specific type of provocation we ought to pursue is one that encourages collective and disruptive struggles against inequality (Forrest 2016). From this perspective, the purpose of examining meaning-making and the contingencies and tensions it entails is not just to

9 Peregrine Schwartz-Shea's and Carolyn Holmes's contributions to this symposium also use the word "provocation" in relation to Taylor's article, but in a somewhat different way than I do. Their essays underscore how Taylor's article directly provoked empirical political scientists to conduct more—or at least better appreciate—interpretive research. My essay emphasizes how, according to Taylor, interpretive research itself should provoke its readers.

generically challenge readers' "intuitions" and expand their sense of sociopolitical possibility. Rather, it is to highlight a specific set of possibilities—namely, possibilities for ordinary people to rework this world in their favor. Scholarship should be provocative in the sense that it pushes readers to better appreciate, explore, and broadcast these possibilities.

One can see this kind of provocation clearly, for instance, in Frances Fox Piven and Richard Cloward's (often) interpretive research about social policy, social movements, and elections in the United States (Piven and Cloward 1971, 1977, 1988). Through their research, Piven and Cloward successfully examined multiple points of contingency underlying the sociopolitical marginalization of poor people, with the explicit aim of locating possibilities to disrupt that marginalization. Most importantly for this discussion, they examined the institutionalized, oppressive, and contingent meanings attached to welfare and work in the US, focusing on how ordinary people could potentially diminish the strength of these meanings and enable more helpful antipoverty policymaking.[10] Embracing this more disruptive and egalitarian type of research may not be the only way to extend and refine Taylor's move toward provocation. But it demonstrates the fruitfulness of making such an extension.

On the topic of political science's analytical means, Taylor's main alternative to "brute data" is, as he says, "text" (i.e., writing and speech) and "text-analogues" (i.e., practices, depictions, built environments, and anything else that communicates to audiences and can, thus, be 'read' as text). Studying "text" and "text-analogues" across different contexts, Taylor argues, is how one gets beyond the world of "brute data" and into the realm of meaning-making. It reveals how societal actors assemble, reproduce, rework, and challenge the meanings—and, thus, the *existence*—of phenomena like negotiation and legitimacy. And, importantly, it allows us to locate and theorize about important sociopolitical contingencies and tensions that arise out of this process.

I, of course, agree with Taylor about the importance of studying "text" and "text-analogues." Indeed, for contemporary interpretive scholars, much of what he says about the importance of all things textual has (to his credit) achieved the status of common sense. However, if *all* one does is examine the operation of "text" and "text-analogues" for different people in different contexts, then any resulting theory of meaning-making and the contingencies that arise out of it will remain incomplete. There is something else we need to examine, which Taylor gestures towards but, in my opinion, does not emphasize nearly enough. That something else is the changing systemic, political-economic relations that surround each textual situation, what urban geographers Neil Brenner, Jamie Peck, and Nik Theodore (2010) call the "context of context."

For years now, empirical and theoretical studies have shown that one cannot follow the twists and turns of meaning-making without attending to such systemic relations.[11] For example, Piven and Cloward (1977) demonstrated that the meanings and tensions surrounding welfare and work in the United States have been significantly shaped by broad political-economic transformations, especially whether those transformations weaken "the structures of daily life" and "the regulatory capacities of these structures" (11). Likewise, Stuart Hall's famous essay on "Race, Articulation, and Societies Structured in Dominance" ([1980] 2021) showed that social scientists cannot explain the construction, meaning, and contingency of race without examining its variegated conjunctions with capitalist political economy.[12] Yes, "texts" and "text-analogues" themselves help give rise to the changing systemic relations referenced by scholars like Piven and Cloward and Hall—an insight that these scholars as well as Taylor are eager to acknowledge. But the inverse, they show, is equally true.[13]

In a few instances, Taylor's article signals the explanatory importance of systemic relations. For example, early on, he makes passing reference to the notion that "hierarchical relations of power and command" grant some textual elements and, consequently, some meanings more societal and political influence than others (1971, 12). Additionally, in discussing the turmoil of the 1960s, he hypothesizes that certain sociopolitically significant meanings may have "gone sour" due to the loss of a political-economic "horizon to be attained by future greater production (as opposed to social transformation)"

10 See also Schram (2002).

11 On this point, see also Natasha Behl's discussion of "embodied positionalities" in this symposium. Her essay at least implies that one the major ways in which systemic, political-economic relations influence meaning-making is by structuring these positionalities across various contexts.

12 As Hall ([1980] 2021, 198-99) aptly states, "Unless one attributes to race a single, unitary transhistorical character – such that wherever and whenever it appears it always assumes the same autonomous features, which can be theoretically explained, perhaps, by some general theory of prejudice in human nature […] – then one must deal with the historical specificity of race in the modern world. Here one is then obliged to agree that race relations are directly linked with economic [i.e., systemic] processes."

13 A more recent qualitative publication that skillfully conveys this mutually constitutive relationship between the systemic and the situational and textual is Fairbanks (2009).

(42). However, outside of these instances, he never really explores how systemic relations do not just grow out of but also, in turn, constitute the specific contexts in which "texts" and "text-analogues" operate and meaning-making occurs. Maybe this was simply because he already had so much to write about. Or maybe he was wary of unintentionally giving ground to overly systemic—especially "classical" or structuralist Marxist—analyses of society.[14] Regardless, if contemporary scholars want to fully realize Taylor's aspirations for an interpretive political science, we need to fill this gap in his argument ourselves—underscoring the explanatory importance of the systemic *as well as* the textual and situational.

My own scholarship would have been inconceivable without extending Taylor's insights in the ways discussed above. For one, my aim in exploring contingencies and tensions in the work of social justice organizing was not to generically challenge readers to "develop" their "intuitions" and expand their sense of possibility; it was, as I stated above, to underscore specific (and often overlooked) possibilities to improve this work and more effectively support disruptive and egalitarian movements of ordinary people. Furthermore, while I closely examined how organizers engaged with "texts" and "text-analogues" across various contexts, to explain why and how they made their world meaningful and influenced its future, I also had to go a little further. I had to examine how (and with what limits) the systemic relations associated with neoliberal capitalism structured their textual engagements. None of this is to say that less avowedly disruptive and egalitarian kinds of provocation or more bounded analyses of "texts" and "text-analogues" are somehow useless. Not at all. What I am saying, however, is that, in considering and building on Taylor's insights, scholars should sometimes also push past those limits.

## Conclusion

I do not have much to add by way of a conclusion, except to just say thanks. Thanks to Charles Taylor for giving us his brilliant article. And thanks to each of the scholars who initially read it, grappled with his ideas, built on his ideas, taught those ideas, and applied them to their research. Without the work of these scholars and all who supported them, the interpretive tradition of qualitative political science would certainly still exist. But it would not be as rich and productive.

## Acknowledgements

Many thanks to Sarah El-Kazaz, Matthew Longo, Peregrine Schwartz-Shea, Dvora Yanow, and an anonymous reviewer for their helpful thoughts regarding this article.

## References

Barrow, Clyde. 2008. "The Intellectual Origins of New Political Science." *New Political Science* 30 (2): 215-44.

Brady, Henry E., and David Collier, eds. 2004. *Rethinking Social Inquiry: Diverse Tools, Shared Standards*. Lanham: Rowman & Littlefield.

Brenner, Neil, Jamie Peck, and Nik Theodore. 2010. "Variegated Neoliberalization: Geographies, Modalities, Pathways." *Global Networks* 10, no. 2 (April): 182-222.

Caldwell, Marc. 2009. "Charles Taylor and the Pre-history of British Cultural Studies." *Critical Arts* 23(3): 342-73.

Fairbanks, Robert P. 2009. *How It Works: Recovering Citizens in Post-Welfare Philadelphia*. Chicago: University of Chicago Press.

Ferguson, James. 1994. *The Anti-Politics Machine: "Development," Depoliticization, and Bureaucratic Power in Lesotho*. Minneapolis: University of Minnesota Press.

Forrest, M. David. 2016. "For a Ruthless Criticism of U.S. Politics." *Polity* 48, no. 1 (January): 5-28.

Forrest, M. David. 2017. "Engaging and Disrupting Power: The Public Value of Political Ethnography." *PS: Political Science & Politics* 50, no. 1 (January): 109-13.

Forrest, David. 2022. *A Voice but No Power: Organizing for Social Justice in Minneapolis*. Minneapolis: University of Minnesota Press.

Fraser, Ian. 2003. "Charles Taylor, Marx and Marxism." *Political Studies* 51, no. 4 (December): 759-74.

Gaventa, John. 1980. *Power and Powerlessness: Quiescence and Rebellion in an Appalachian Valley*. Urbana: University of Illinois Press.

George, Alexander L., and Andrew Bennett. 2004. *Case Studies and Theory Development in the Social Sciences*. Cambridge: MIT Press.

Hall, Stuart. (1980) 2021. "Race, Articulation, and Societies Structured in Dominance." In *Selected Writing on Race and Difference*, edited by Paul Gilroy and Ruth Wilson Gilmore,195-245. Durham: Duke University Press.

14 Taylor, we should remember, came of age intellectually and politically as a participant in the early days of the British New Left, at Oxford University. In 1957, along with Stuart Hall and a couple other students, he founded and edited the *Universities and Left Review*, one of the progenitors of the *New Left Review*. Many in this group were wary (in Taylor's case, perhaps too wary) of structuralist or "classical" Marxist analyses, which they associated with the moral and practical failures of Stalinism (Caldwell 2009). This wariness certainly carried over into his years as a professor (Fraser 2003).

Lupia, Arthur. 2014. "What is the Value of Social Science? Challenges for Researchers and Government Funders." *PS: Political Science and Politics* 47, no. 1 (January): 1-7.

McCoy, Charles A., and John Playford, eds. 1967. *Apolitical Politics: A Critique of Behavioralism*. New York: Crowell.

Piven, Frances Fox and Richard A. Cloward. 1971. *Regulating the Poor: The Functions of Public Welfare*. New York: Vintage.

Piven, Frances Fox and Richard A. Cloward. 1977. *Poor People's Movements: Why They Succeed, How They Fail.* New York: Vintage.

Piven, Frances Fox and Richard A. Cloward. 1988. *Why Americans Don't Vote*. New York: Pantheon.

Polsby, Nelson W. 1960. "How to Study Community Power: The Pluralist Alternative." *Journal of Politics* 22, no. 3 (August): 474-84.

Schatz, Edward, ed. 2009. *Political Ethnography: What Immersion Contributes to the Study of Power*. Chicago: University of Chicago Press.

Schram, Sanford F. 2002. *Praxis for the Poor: Piven and Cloward and the Future of Social Science in Social Welfare*. New York: New York University Press.

Taylor, Charles. 1971. "Interpretation and the Sciences of Man." *Review of Metaphysics* 25, no. 1 (September): 3-51.

Wolin, Sheldon. 1969. "Political Theory as a Vocation." *American Political Science Review* 63, no. 4 (December): 1062-82.

Qualitative and Multi-Method Research Fall 2023, Volume 21.2 https://doi.org/10.5281/zenodo.8326504

# Critical Rereading of Charles Taylor and Reflections on Interpretive Research in Political Science

Natasha Behl
*Arizona State University*

Charles Taylor (1971, 3) asked if interpretation is essential to social scientific explanation. In answering this question, he made (and continues to make) multiple radical contributions to political science, which include challenging understandings of science based on verification, presumptions about the universality of Western ideas, and explanations of political behavior centered on individualist assumptions. In doing so, Taylor opened up the possibility for interpretivism within political science by calling for a kind of social scientific explanation that centers the meaning making of social actors. His contributions helped to shape the current contours of interpretivism in political science. Whereas there is no single definition of interpretivism in the social sciences, interpretive research does share some characteristics, including (1) an understanding of knowledge as "historically situated" and "entangled in power relationships"; (2) an understanding of the world as "socially made"; (3) a skepticism of "individualist assumptions" that dominate rational-choice and behavioralist approaches; and (4) an interest in "culture" and "language" (Wedeen 2009, 80-2).

Taylor's insights continue to have a lasting impact on the discipline of political science by opening up alternative ways of doing empirical scholarship, but alone they are insufficient for scholarship today. Interpretivism has created a space, perhaps a home, for those of us who are marginalized epistemologically and methodologically within political science. However, interpretivism has not necessarily created a space for those of us who also find ourselves marginalized due to our embodied positionality. Taylor made this leap possible for interpretivism, a leap that scholars can take today through sustained dialogue with feminist, race, and decolonial scholars. Such a dialogue could enable scholars to focus on the gendered, racialized, and colonial logics of politics as a practice and political science as a discipline. It could also enable scholars to become attentive to the "twin battle over the politics of knowledge" that feminist, race, and decolonial scholars have fought and continue to fight—first, to be recognized in academic communities, and second, to have their arguments included as legitimate forms of knowledge (Ackerly 2021, 402; see also Smith 2012; Simpson 2014; Collins 2019).

In this essay, I engage in a critical rereading of Charles Taylor's 1971 article by situating it within current interpretive methodological thinking, reading it in dialogue with a rich but largely overlooked lineage of feminist, intersectional, race, and decolonial scholars, and using it as a lens to reflect more broadly on political science as a discipline and the interpretive research community within it.[1] I demonstrate how I came to center questions of the politics of knowledge as they relate to non-normative bodies and identities through a sustained engagement with both interpretivism and alternative academic histories rooted in feminist, intersectional, race, and decolonial scholarship, which already exist but have been marginalized by mainstream social science through active erasure and whitening (Jordan-Zachery 2007; Alexander-Floyd 2012; Tuck and Yang 2012; Bilge 2013a, 2013b) and deliberate disregard and exclusion (Hawkesworth 2005; Tickner 2015; Ahrens et al. 2018).

## Epistemic Oppression, Injustice, and Violence

In my writing I speak, at times, of my discomfort with social scientific categories because they often result in a kind of research that feels "cold" and "disconnected" from the pain, injustice, and violence that animates my research (Behl 2019b, 91; see also Doty 2004, 378; Isoke 2018, 163). Part of my discomfort came initially from this vague sense that perhaps these categories are perpetuating the very injustices I am trying to understand and eradicate (Behl 2019b, 91). I often struggled to find the language to name epistemic injustices alongside racial and gendered violence that are often the norm for many of us who find ourselves at the margins of academia due to our epistemological and methodological choices and our embodied positionalities. I often thought that this discomfort was a deficiency in me, an inadequacy in my way of thinking.

What I never considered was that my "vague sense" that followed me in my undergraduate and graduate training was perhaps a kind of fledgling "intuition" that I have developed and cultivated with time and self-reflection. As Taylor (1971, 51) explained, interpretivism is "founded on intuitions" that require "a high degree of self-knowledge…[because our capacity] to be understood is rooted in our own self-definitions, hence in what we are." This is one of Taylor's (1971, 51) radical interventions which was and continues to be "shocking…to…modern science." He called for political science "to go beyond the bounds of a science based on verification to one which would study the intersubjective and common meanings embedded in social reality" (45).

Another radical intervention was Taylor's (1971, 33) critique of mainstream comparative politics, which he explained requires a "universal vocabulary of behavior" to compare different "practices of different societies in the same conceptual web." The not surprising result, according to Taylor (34), "is a theory of political development which places the Atlantic-type polity at the summit of human political achievement." For Taylor (40), "The inability to recognize…intersubjective meanings is…inseparably linked with the belief in the universality of North Atlantic behavior types."

Yet another key intervention was Taylor's (45) explanation of why political science fails to come "to grips with important problems of our day." For Taylor (32) there is no "place in mainstream social science" that can account for an "I" that can also be a "we." He explained that the exclusion of the possibility of the "communal" comes from the influence of an "epistemological tradition for which all knowledge has to be reconstructed from the impressions imprinted on the individual subject" (32).

What I find so compelling about Taylor's article is its ability to name the epistemic discomfort and injustice that I have struggled to name, make sense of, and navigate in political science. He explained why it is so difficult to navigate this terrain as an empirical scholar informed by interpretive methodologies. He argued that "intersubjective [meanings] are constitutive of [social] reality" (30). However, intersubjective meanings "fall through the net of mainstream social science" (31). These kinds of common meanings "can find no place in its categories" (31).

These aspects of Taylor's article resonate deeply with my own scholarship and academic journey as Taylor identified what I so often find myself writing against—the presumed universality of Western ideas, political secularism, and political development. In my earlier research, I tried to make sense of the lived experiences of minority Sikh women in Punjab, India through the established and rehearsed categories of political science. But these categories failed to fully reflect the fullness of these women's actions and meanings, especially as it relates to their interconnected understandings of human and divine agency and their contingent alignment of spiritual and gender-based liberation (Behl 2019a). I centered the lived experiences of non-Western women and non-secular women to demonstrate how secular mechanisms designed for inclusion can exclude, while forms of devotion assumed to be undemocratic can be inclusionary (Behl 2019a).

More recently, I have tried to make sense of the protesting farmers in India through the theories of rational collective action. But the farmers could not be reduced to the individualist assumptions that characterize much social movement scholarship (Behl 2022). So much was overlooked and lost through an imposition of assumptions about their individualized motives

1 Taylor himself did not use either "positionality" or "reflexivity" in his 1971 article as they were not yet part of academic discourse, whether in the methods literature or in discussions of methodology (as Yanow and Schwartz-Shea note in their essays, this symposium).

implicit in these theories and through an "imperialism of categories" designed to capture and explain the other (Rudolph 2005; Smith 2012). Social scientific categories failed to fully reflect the protestors' collective recasting of authoritarian democracy towards more inclusive and egalitarian democratic practices. I centered the political actions of non-Western farmers and laborers to show how an embrace of religion is not necessarily antithetical to liberal democracy but might serve to protect it (Behl 2022).

In my own academic journey, I coupled interpretivism with the insights of feminist, intersectional, race, and decolonial scholars to ask, how might scholars bring an epistemic humility to the research process (Bierria 2020)? How might our research questions, accepted beliefs, and shared assumptions transform if the very racialized and gendered individuals who are so often the objects of social scientific inquiry are understood as knowing subjects, as critical social theorists, and as democratic theorists and practitioners (Collins 2019; Pineda 2021)? How might the form and content of scholarship transform if scholars acknowledge that theorizing is not limited to elite university spaces but also happens within families, religious spaces, social movements, and communities (Simpson 2014)?

Feminist, intersectional, race, and decolonial scholars bring a sustained focus to the question of how epistemic power intersects with non-normative bodies and identities to create "objective" knowers and "valid" knowledge (Bonilla-Silva and Zuberi 2008; Smith 2012; Bierria 2020; Ackerly 2021). In doing so, these scholars provide insight on how to map epistemic privilege and oppression in the research process and how to locate potential sources to challenge it. These insights open up the possibility of new theoretical and empirical horizons by challenging assumptions about who is considered a legitimate knowledge producer, what counts as valid forms of knowing, what does it mean to theorize, and where does the labor of theorizing occur (Simpson 2014, 7; Collins 2019, 10; Bierria 2020, 301). Through a sustained dialogue with these communities of scholars, I learned to name epistemic oppressions, cultivate my "intuition," and preserve my epistemic resistance even as it was being disciplined out of me, even as it was being snuffed in the name of objectivity, validity, and replicability.

What I held on to at these moments of epistemic violence were Gloria Anzaldúa's (1981, 168-9) words from her letter to other third world women writers:

> I must keep the spirit of my revolt and myself alive. Because the world I create in the writing compensates for what the real world does not give me. By writing I put order in the world, give it a handle so I can grasp it. I write because life does not appease my appetites and hungers. I write to record what others erase when I speak, to rewrite the stories others have miswritten about me, about you. To become more intimate with myself and you. To discover myself, to preserve myself, to make myself, to achieve autonomy.

I sustained myself and my epistemic resistance in a hostile discipline by returning to the words of women of color scholars and third world feminists, like Gloria Anzaldúa, because their words helped me understand that the deficiency is not in me, nor is it in my thinking. Rather the deficiencies are in racialized and gendered ways of knowing, the inadequacies are in colonized methodologies that masquerade as objective and neutral. In these moments, I followed Gloria Anzaldúa (1981, 173), who calls on us to "throw away abstraction" and "Write with your eyes like painters, with your ears like musicians, with your feet like dancers. You are the truthsayer with quill and torch. Write with your tongues of fire. Don't let the pen banish you from yourself. Don't let the ink coagulate in your pens. Don't let the censors snuff out the spark, nor the gags muffle your voice. Put your shit on the paper."

## Conclusion

The impact of Taylor's article on the discipline is unquestionable given that now some fifty years after its publication, we can collectively speak of an "interpretive turn" in political science where some scholars are engaged in interpretive empirical scholarship (Yanow and Schwartz-Shea 2006; Brodkin 2017). Some scholars celebrate interpretivism for its ability to question prevailing paradigms while inviting "novel ways of imagining the political" (Wedeen 2009; see also Pachirat 2009). While others champion interpretivism because it opens up the possibility of "theoretical vibrancy" and "epistemological innovation" (Schatz 2009). I, like so many others, see promise in interpretivism, especially in its ability to "disrupt forms of power" (Forrest 2017). I, like so many others, am grateful for Taylor's contributions to and insights on political science. For these reasons and many more, interpretivism has become a home, a place of belonging for many of us whose epistemologies and methodologies are deemed unscientific and invalid by the dominant norms of political science.

As Robin Turner (2022) explains, some of us have found a "pathway back" to political science "via interpretive political science…(despite its whiteness)." And yet, this space of interpretivism can be and has been uncomfortable and painful for those of us whose non-normative epistemologies and methodologies intersect with our non-normative bodies and identities (Brown 2007; Alexander-Floyd 2015; Fujji 2016; Behl 2017, 2019b; Brown 2019; Jordan-Zachery 2019; Willoughby-Herard 2019; Turner 2022).

Taylor sought to decenter universality, individualist assumptions, and science in political science. This radical intervention was and continues to be necessary, but on its own it is insufficient. As Lee Ann Fujii (2016) reminds us, we must call out ways of seeing and explaining the world that are rooted in a "racialized lens of whiteness," which is simultaneously assumed to be "neutral, unraced, and ungendered and therefore 'scientifically' sound." I fear that without a more sustained focus on epistemic privilege and oppression, political science as a discipline, and perhaps the interpretive research community within it, may not fully confront "the racist origins of American political science" (McClain 2021, 7) and may not strive to construct a more just future.

## Acknowledgements

An earlier version of this essay was presented in a roundtable on Taylor's 1971 article held at the 2022 WPSA conference in Portland, OR, sponsored by the Interpretation and Methods Section. I am grateful to Dvora Yanow, Peri Schwartz-Shea, and Robin Turner for encouraging me to pursue this line of inquiry and for their insightful thoughts on previous drafts.

## References

Ackerly, Brooke. 2021. "Epistemic Oppression, Grounded Normative Theory, and an 'Ethnographic Sensibility.'" *Contemporary Political Theory* 20, no. 2 (June): 385-418.

Ahrens, Petra, Karen Celis, Sarah Childs, Isabelle Engeli, Elizabeth Evans, and Liza Mugge. 2018. "Politics and Gender: Rocking Political Science and Creating New Horizons." *European Journal of Politics and Gender* 1 (1–2): 3–16.

Alexander-Floyd, Nikol. 2012. "Disappearing Acts: Reclaiming Intersectionality in the Social Sciences in a Post-Black Feminist Era." *Feminist Formations* 24, no. 1 (Spring): 1–25.

Alexander-Floyd, Nikol. 2015. "Women of Color, Space Invaders, and Political Science: Practical Strategies for Transforming Institutional Practices." *PS: Political Science & Politics* 48, no. 3 (July) :464-68.

Anzaldúa, Gloria. 1981. "Speaking in Tongues: A Letter to Third World Women Writers." In *This*

*Bridge Called My Back: Writings by Radical Women of Color*, edited by Gloria Anzaldúa and Cherrie Moraga, 165–74. New York:Kitchen Table/Women of Color.

Behl, Natasha. 2017. "Diasporic Researcher: An Autoethnographic Analysis of Gender and Race in Political Science." *Politics, Groups, and Identities* 5 (4): 580–98.

Behl, Natasha. 2019a. *Gendered Citizenship: Understanding Gendered Violence in Democratic India*. New York: Oxford University Press.

Behl, Natasha. 2019b. Mapping Movements and Motivations: An Autoethnographic Analysisof Racial, Gendered, and Epistemic Violence in Academia. *Feminist Formations* 31, no. 1 (Spring): 85–102.

Behl, Natasha. 2022. "India's Farmers' Protest: An Inclusive Vision of Indian Democracy." *American Political Science Review* 116, no. 3 (August): 1141–46.

Bierria, Alisa. 2020. "Grassroots Philosophy and Going Against the Grain." In *A Guide to Field Philosophy: Case Studies and Practical Strategies*, edited by Evelyn Brister and Robert Frodeman, 301-17. New York: Routledge Press.

Bilge, Sirma. 2013a. "Intersectionality Undone: Saving Intersectionality from Feminist Intersectionality Studies." *Du Bois Review* 10, no. 2 (Fall): 405–24.

Bilge, Sirma. 2013b. "Whitening Intersectionality: Evanescence of Race in Intersectionality Scholarship." In Racism and Sociology, edited by Wulf D. Hund and Alana Lentin, 175– 206. Berlin: Lit Verlag.

Bonilla-Silva, Eduardo, and Tukufu Zuberi. 2008. *White Logic, White Methods: Racism and Methodology*. Lanham: Rowman & Littlefield.

Brodkin, Evelyn Z. 2017. "The Ethnographic Turn in Political Science: Reflections on the State of the Art." *PS: Political Science & Politics* 50, no. 1 (January): 131–4.

Brown, Nadia. 2019. "Mentoring, Sexual Harassment, and Black Women Academics." *Journal of Women, Politics & Policy* 40 (1): 166-73.

Brown, Ruth Nicole. 2007. "Persephone's Triumph: Reflection of a Young Black Woman Becoming a Real Political Scientist." *Qualitative Inquiry* 13, no. 5 (July): 650–59.

Collins, Patricia Hill. 2019. *Intersectionality as Critical Social Theory*. Durham: Duke University Press.

Doty, Roxanne Lynn. 2004. "Maladies of Our Souls: Identity and Voice in the Writing of Academic International Relations."

*Cambridge Review of International Affairs* 17 (2): 377–92.

Forrest, M. David. 2017. "Engaging and Disrupting Power: The Public Value of Political Ethnography." *PS: Political Science & Politics* 50, no. 1 (January): 109–13.

Fujii, Lee Ann. 2016. "Changing Disciplines" (keynote address, International Studies Association- Northeast, Baltimore, MD, November 4).

Hawkesworth, Mary. 2005. "Engendering Political Science: An Immodest Proposal." *Politics & Gender* 1, no. 1 (March): 141–56.

Isoke, Zenzele. 2018. "Black Ethnography, Black(Female)Aesthetics: Thinking/Writing/Saying/Sounding Black Political Life." *Theory & Event* 21, no. 1 (January): 148–68.

Jordan-Zachery, Julia. 2007. "Am I a Black Woman or a Woman Who Is Black? A Few Thoughts on the Meaning of Intersectionality." *Politics & Gender* 3, no. 2 (June): 254–63.

Jordan-Zachery, Julia. 2019. "Licking Salt: A Black Woman's Tale of Betrayal, Adversity, and Survival." *Feminist Formations* 31, no. 1 (Spring):67-84.

McClain, Paula D. 2021. "Crises, Race, Acknowledgement: The Centrality of Race, Ethnicity, and Politics to the Future of Political Science." *Perspectives on Politics* 19, no. 1 (March):7-18.

Pachirat, Timothy. 2009. "The Political in Political Ethnography: Dispatches from the Kill Floor." In *Political Ethnography: What Immersion Contributes to the Study of Power*, edited by Edward Schatz, 143–62. Chicago: University of Chicago Press.

Pineda, Erin. 2021. *Seeing Like an Activist.* Oxford: Oxford University Press.

Rudolph, Susanne. 2005. "Imperialism of Categories: Situating Knowledge in a Globalizing World." *Perspectives on Politics* 3, no. 1 (March): 5– 14.

Schatz, Edward. 2009. "Introduction: Ethnographic Immersion and the Study of Politics." In *Political Ethnography*, edited by Edward Schatz, 1– 22. Chicago: University of Chicago Press.

Simpson, Leanne. 2014. "Land as Pedagogy: Nishnaabeg Intelligence and Rebellious Transformation." *Decolonization: Indigeneity, Education & Society* 3 (3): 1-25.

Smith, Linda Tuhiwai. 2012. *Decolonizing Methodologies: Research and Indigenous Peoples*. London: Zed Books Ltd.

Taylor, Charles. 1971. "Interpretation and the Sciences of Man." *The Review of Metaphysics* 25, no. 1 (September): 3-51.

Tickner, J. Ann. 2015. "Revisiting IR in a Time of Crisis." *International Feminist Journal of Politics* 17 (4): 536–53.

Tuck, Eve, and K. Wayne Yang. 2012. "Decolonization is Not a Metaphor," *Decolonization: Indigeneity,Education & Society* 1 (1): 1-40.

Turner, Robin L. 2022. "Reflections on Multiple Consciousness and Space Invasion in Comparative Political Science." Paper presented at the Annual Meeting of the American Political Science Association, Montreal, Canada, September 15.

Wedeen, Lisa. 2009. "Ethnography as Interpretive Enterprise." In *Political Ethnography: What Immersion Contributes to the Study of Power*, edited by Edward Schatz, 75-93. Chicago: University of Chicago Press.

Willoughby-Herard, Tiffany. 2019. "Poetic Labors and Challenging Political Science: An Epistolary Poem." *Journal of Women, Politics & Policy* 40 (1): 228-35.

Yanow, Dvora, and Peregrine Schwartz-Shea. 2006. *Interpretation and Method: Empirical Research Methods and the Interpretive Turn*. Armonk: M.E. Sharpe.

# Finding the Bridge: Charles Taylor, Interpretive Methods, and Political Philosophy

Matthew Longo
*Leiden University*

We all have our intellectual debts; Charles Taylor is the source of many of mine. He helped forge a path for interpretive social science I still travel and spelled out the philosophical groundwork by which it might be sustained.[1] My point here is not to revisit his canon, but to ask instead how Taylor's writing on interpretation connects to contemporary debates in political philosophy. As a member of a niche but growing community of political theorists who do interpretivist fieldwork,[2] the question is a pressing one. As a group, we speak a lot about the "ethnographic sensibility"[3] and weave freely between empirical and theoretical idioms. Having done fieldwork, we know a lot about thick description and contextualization, about positionality and perspective. But how exactly can we bridge the divide between these disparate fields? Why might political philosophers benefit from adopting an interpretivist lens? And how do we explain this contrapuntal positioning to skeptics who doubt the value of such work? I think Taylor helps furnish the answer.

In what follows, I try to explain why, using two texts: Taylor's landmark article "Interpretation and the Sciences of Man" (1971), which makes the case for interpretive social science; and "What's Wrong with Negative Liberty" (1979), which situates interpretation at the heart of philosophical debates about freedom. In these works, Taylor shows not merely how interpretivism might be of value to political philosophy, but why it is essential to it. The goal of this paper is to detail how he does so, in the hopes that future scholars interested in the intersection of these fields might profit from his insights in the ways I have.

## Interpretation and the "Meaning of Meaning"

In a short article like this one, there is no space to go into the full defense of interpretive social science that Taylor makes in "Interpretation and the Sciences of Man." Instead, I focus on three specific components of his argument that lay the groundwork for the case I will later make about political philosophy: hermeneutical science; the "meaning" of meaning; and the determinants of evaluation.

First, hermeneutics. All social observations operate with a set of assumptions (stated or unstated) about the world – what it looks like, what we can really know about it, and so forth. The more ordered you imagine the social world to be, the more rule-guided and systematic, the more likely you are to believe that things can be measured (and that future events can be predicted). In Taylor's terminology, under this view the world is comprised of "brute data"[4] —that is, data "whose validity cannot be questioned by offering another interpretation… whose credibility cannot be founded or undermined by further reasoning" (1971, 8). This is the hallmark of "positivism." By contrast, the less rule-guided you think the world is, the less likely that social facts can be received uncomplicatedly. Interpretive methods — Taylor's "hermeneutical science" —take as their object of study something "which in some way is confused, incomplete, cloudy, seemingly contradictory in one way or another, unclear" (1971, 3). The aim of interpretation is to "bring to light an underlying coherence or sense" (1971, 3).

But how do you establish the validity of this starting point as the legitimating core of one's study? Taylor argues that a division must be wrought between "meaning and expression," between the sense that undergirds a practice or an utterance and its manifestation in the world. The goal of a hermeneutical science, he argues, must be to make coherent sense of this meaning; in so doing, it also reveals the implausibility of a mainstream social science that simply pursues "brute data" and aims to establish facts beyond reproach or critique. In such a rendering, meaning and expression are collapsed together. This, Taylor contends, is inherently reductive; it can only be

1 For a broad treatment of this field, both in its empirical and theoretical forms, see Yanow and Schwartz-Shea 2014; it is also spelled out in Dvora Yanow's contribution to this symposium.

2 Examples of recent works include Iqtidar 2011, Zacka 2017, Longo 2018, Behl 2019, and Blajer de la Garza (n.d.).

3 See e.g., Schatz 2009; for a theoretical treatment of this kind of work, see Herzog and Zacka 2017, Longo and Zacka 2019, and Zacka et al 2020.

4 For a deeper engagement of this idea of brute data, see the contributions by Peregrine Schwartz-Shea and David Forrest in this symposium.

achieved by filtering away the complexity of the world that it is trying to explain.

A second point regards the nature of meaning itself —that is, the "meaning" of meaning. Those of us who work in the field of interpretation use the word "meaning" a lot, but we don't always take the time to spell out what our usage entails.[5] Taylor (1971) provides a considered definition.

> When we speak of the "meaning" of a given predicament, we are using a concept which has the following articulation. (1) Meaning is for a subject: it is not the meaning of the situation in vacuo … (2) Meaning is of something; that is, we can distinguish between a given element – situation, action, or whatever – and its meaning … (3) things only have meaning in a field, that is, in relation to the meanings of other things … Meaning in this sense – let us call it experiential meaning – thus is for a subject, of something, in a field. (11-2)

This definition of meaning highlights the role of the subject, and subjectivity, in meaning-making, which take place against a "background of desire, feeling, emotion" (1971, 13) and renders untenable any research that severs the observer from the object of study. Taylor highlights this feature again in his discussion of self-interpretation. "There is no such thing as the structure of meanings for [man] independently of his interpretation of them," Taylor writes, "for one is woven into the other" (1971, 16). Thinking about meaning as experiential and positional allows Taylor to flesh out the way that meaning embeds itself in communities and forms of collective understanding—via inter-subjectivity. It also runs counter to studies built upon individualist models of agency that provide the basis for most empirical social science.

> What the ontology of mainstream social science lacks is the notion of meaning as not simply for an individual subject; of a subject who can be a "we" as well as an "I." The exclusion of this possibility, of the communal, comes once again from the baleful influence of the epistemological tradition for which all knowledge has to be reconstructed from the impressions imprinted on the individual subject. (1971, 32)

This problem of I-centrism, of placing the self-contained individual at the heart of analysis, returns in Taylor's discussion of political philosophy.[6]

The final point I wish to highlight pertains to evaluation. Part of his critique of empiricist political science is that it fails to grasp meanings that are value laden. He discusses this via the concept of "legitimacy," which by its nature includes attributes that are moral or evaluative. Consequently, it is a concept whose broad contours cannot be satisfactorily defined or measured by positivist political science, which can assess whether a given population considers a specific regime or policy to be legitimate (via attitudinal measures, for example), but can offer no reflection about legitimacy *as such*. As a result of this inability to negotiate evaluative claims, mainstream political scientists cannot grapple with the kinds of signification their claims embody. Additionally, in doing so, they suggest a kind of universality—or objectivity—in the (unexamined) meaning of the concepts they use, which tends to generate a western-centric (and in other cases, hetero-normative) bias in their usage. This discussion of values segues seamlessly into the next question, about whether and how interpretive research might contribute to political philosophy—discussed below with regards to the problem of freedom.

## Philosophy and "Desires about Desires"

In "What's Wrong with Negative Liberty," Taylor responds to Isaiah Berlin's foundational treatment of the problem of freedom in his essay, "Two Concepts of Liberty," originally given as a lecture in 1958, in which Berlin (1997) makes a distinction between positive and negative liberty. For our purposes, we can define the two terms roughly as follows. Negative liberty refers to freedom from interference, the area around the self that other actors cannot enter, in which we are left to do as we please. This is frequently referred to as "freedom from," or *independence*. By contrast, positive liberty is about self-mastery, the ability to determine who we are. This is often referred to as "freedom to," or *autonomy*. Taylor's point isn't that there is anything inherently wrong with this division, but rather that with time these positions have become caricatures—with a vague and expansive positive liberty (that may go so far as to justify totalitarianism) on the one hand, and a tight and parsimonious negative liberty (too narrow to be useful) on the other.[7] This caricature obscures the real nexus of contestation, which, Taylor contends, isn't about liberty at all but about interpretation.

To reclaim the debate over liberty from its state of caricature, Taylor re-conceptualizes positive/negative

5 Exceptions to this are found in the edited volumes by Schatz (2009) and Yanow and Schwartz-Shea (2014); see for example contributions by Pachirat (2009), Dow (2014), Yanow (2014), and Soss (2014).

6 Individual agency models remain the norm in political philosophy, although there are notable exceptions. For a critique see Sandel (1982).

7 These caricatures remain familiar. Notably, theories of republican freedom avoid this pitfall (Pettit 1997; Skinner 1998).

liberty as a distinction between what he calls an "exercise concept" and an "opportunity concept." What Berlin calls positive liberty, Taylor argues, is really about "exercising control over one's life"—one is free insofar as they have "effectively determined oneself and the shape of one's life" ([1979] 2006, 143). By contrast, negative liberty is "an opportunity-concept, where being free is a matter of what we can do, of what it is open to us to do, whether or not we do anything to exercise these options" ([1979] 2006, 143). By re-working the terms in this way, Taylor alerts us to the problem: exercise is a messy, heavily subjective concept; opportunity is comparatively easy to define (or, you might say, "brute"). Hence the caricatures: positive liberty comes to mean something expansive that could morph into anything (even totalitarianism); negative liberty shrinks into the tightest carapace, which is defensible, but too delimited to have much value.

With this re-working, Taylor can point us to what he believes is the source of the problem: namely that in preferring the opportunity concept over the exercise concept, most contemporary political theorists have eliminated much of what is important (and meaningful) about the problem of freedom, as at the core of exercise concepts are battles over values—a conceptual terrain in which one must "fight to discriminate the good from the bad … fight, for instance, for a view of individual self-realization against various notions of collective self-realization, of a nation, or a class" ([1979] 2006, 145). Absent such valuations, we are left with what he calls the "Maginot Line" theory of freedom—its caricatured form, the simple absence of external constraints—which is overly minimalistic and privileges a reductive account of the world, contra the complexity that is its core. This is precisely the mirror of his critique in "Interpretation and the Sciences of Man" (1971) of a social science that mainly pursues "brute data" beyond interpretation or reproach.

The Maginot Line theory, Taylor contends, prevents us from making judgments based on meaningfulness – about the "meaning of meaning," you might say. The now-famous example he gives is of traffic lights. Clearly traffic lights present restrictions on freedom: when the light is red, we are not permitted to proceed forward. But none of us consider these impediments to be *meaningful*. By contrast, we care deeply about other restrictions, such as on religious freedom. Taylor's point is not to adjudicate the validity of these views, but rather to point out that narrow theories—especially those that quantify objects of study (that turn them into something "brute") —cannot accommodate this judgment. "There are discriminations to be made," he writes, "[But the Maginot Line theory] has no place for the notion of significance. It will allow only for purely quantitative judgments" ([1979] 2006,150). Absent such judgment, he continues, we would have no way of distinguishing between the freedoms of Britain (religious freedom; many traffic lights) with that of Albania (no religious freedom; minimal traffic lights). What we need is a theory that helps us distinguish why some freedoms are more meaningful than others. This brings us back, of course, to *interpretation*—that is, the ways in which meaning is "for a subject, of something, in a field."

Taylor argues that a theory of freedom must begin instead with what he calls "strong evaluation" —the fact that some things matter more to us than others and that these judgments are essential to our self-identity as people. To make this point he cleaves a distinction between our desires and what he calls our "desires about desires."

> We human subjects are not only subjects of first-order desires, but of second-order desires, desires about desires. We experience our desires and purposes as qualitatively discriminated, as higher or lower, noble or base, integrated or fragmented, significant or trivial, good and bad. ([1979] 2006, 152)

In other words, we don't just want to make choices, we want to make good choices – the kind that make us happy or proud. We don't just want to be free to act, but to act well, however it is that we define this term.

Once we understand the nature of these second-order desires, we realize that obstacles to freedom are also internal—they are *intra*-subjective, running counter to the unitary (I-centric) agency models that dominate philosophy as much as social science. Therefore, to understand freedom, we need to appreciate the fact that we discriminate between and among our own emotions. Some are "import-attributing" and thus essential to understanding how any of us, individually, come to believe ourselves to be free. For Taylor, to make freedom meaningful there must be a way to incorporate these kinds of second-order desires into our notion of freedom. To do so, we need to escape the language of freedom as a purely opportunity-concept. There must be space for "strong evaluation" —or, we might say, interpretation.

## Conclusion

The aim of this paper was to show how the writings of Charles Taylor can help forge the bridge between interpretive methods and political philosophy. It did so by reading "Interpretation and the Sciences of Man" (1971) and "What's Wrong with Negative Liberty?" (1979) in tandem. In the first article, Taylor shows the problem of reductiveness in empirical social science, how the reliance on "brute data" prevents us from studying the more complex aspects of the social world, and forces us to collapse meaning and expression, rather than appreciate their difference. This produces an impoverished social science. But as we can see from the second article, the same critique can be leveled against much political philosophy. This literature also tends to shy away from meaning (with its subjective and intersubjective connotations) in favor of arguments deemed objective, even quantitative.[8] At the root of both social science and philosophy, then, is the same reductive intuition, and the eschewal of complicated aspects of the world we are trying to understand.

For scholars interested in using interpretive methods to enter debates in political philosophy, Taylor shows the way. Interpretive methods dig into the overlapping dimensions inherent in our positioning in the social world. Political philosophers might prefer to stay aloof on matters of inter- or intra-subjectivity, contextuality, positionality, and so on, but Taylor forces us to ask whether that's possible. He clearly shows the problem of a notion of freedom in which all evaluative complexity has been excised. In doing so he furnishes the ground for scholars to ask new questions of these debates – what *does* make freedom meaningful and how would we know? What is the cost of taking the self-contained individual to be the object of our studies? Interpretive methods help us answer these questions and others like them. Such research generates thickened concepts and layered empirical portrayals that escape—and trouble—overly parsimonious definitions and simplified agency models.

Social science and philosophy go hand in hand; they are in important ways co-constitutive. And interpretation lies at the center of each. It is interpretation that allows for a robust—and evaluative—social science. It also facilitates a more nuanced, thickened, ethics. As someone situated between these fields, I'm grateful to Charles Taylor for giving me the language to appreciate this dynamic, and intellectual resources sufficient to explore it. The aim of this article was to pay this sensibility forward, in the hopes that other researchers might feel the same confidence and enthusiasm charting this course as I have.

## Acknowledgments

I would like to thank the anonymous reviewer for their feedback on this submission, as well as helpful comments on earlier drafts provided by David Forrest, Dvora Yanow, and Peregrine Schwartz-Shea.

## References

Behl, Natasha. 2019. *Gendered Citizenship: Understanding Gendered Violence in Democratic India*. Oxford: Oxford University Press.

Berlin, Isaiah. 1997. "Two Concepts of Liberty." In *Isaiah Berlin: The Proper Study of Mankind*, edited by Henry Hardy and Roger Hausheer, 191-242. New York: Farrar, Straus and Giroux.

Blajer de la Garza, Yuna. n.d. "A House is Not a Home: Citizenship and Belonging in Contemporary Democracies" (book manuscript in progress).

Blau, Adrian. 2017. *Methods in Analytical Political Theory*. Cambridge: Cambridge University Press.

Dow, Douglas C. 2014. "Talking Our Way to Meaningful Explanations," in *Interpretation and Method: Empirical Research Methods and the Interpretive Turn,* edited by Dvora Yanow and Peregrine Schwartz-Shea, 64-79. London: Routledge.

Herzog, Lisa. and Bernardo Zacka. 2017. "Fieldwork in Political Theory: Five Arguments for an Ethnographic Sensibility." *British Journal of Political Science* 49, no. 2 (April): 763-84.

Iqtidar, Humeira. 2011. *Secularizing Islamists? Jama'at-e-Islami and Jama'at-ud-Da'wa in Urban Pakistan*. Chicago: University of Chicago Press.

Leopold, David, and Marc Stears. 2008. *Political Theory: Methods and Approaches*. Oxford: Oxford University Press.

List, Christian and Laura Valentini. 2016. "The Methodology of Political Theory," in *The Oxford Handbook of Philosophical Methodology*, edited by Herman Cappelen, Tamar Szabó Gendler, and John Hawthorne, 525-58. Oxford: Oxford University Press.

Longo, Matthew. 2018. *The Politics of Borders: Sovereignty, Security, and the Citizen after 9/11*. Cambridge: Cambridge University Press.

Longo, Matthew and Bernardo Zacka. 2019. "Political Theory in an Ethnographic Key." *American Political Science Review* 113, no. 4 (November): 1066-70.

Pachirat, Timothy. 2009. "The Political in Political Ethnography: Dispatches from the Kill Floor," in *Political Ethnography: What Immersion Contributes to the Study of Power*, edited by Edward Schatz, 143-62. Chicago: University of Chicago Press.

Pettit, Philip. 1997. *Republicanism: A Theory of Freedom and Government.* Oxford: Oxford University Press.

Sandel, Michael. 1982. *Liberalism and the Limits of Justice.* Cambridge: Cambridge University Press.

8 As evidence for this one can look to recent books and articles about political theory and method and methodology, in which interpretation is barely discussed (Leopold and Stears 2012; List and Valentini 2016; Blau 2017).

Schatz, Edward. 2009. *Political Ethnography: What Immersion Contributes to the Study of Power*. Chicago: University of Chicago Press.

Shapiro, Ian. 2002. "Problems, Methods and Theories in the Study of Politics, or What's Wrong with Political Science and What to Do About It." *Political Theory* 30, no. 4 (August): 596-619.

Skinner, Quentin. 1998. *Liberty Before Liberalism*. Cambridge: Cambridge University Press.

Soss, Joe. 2014. "Talking Our Way to Meaningful Explanations." In *Interpretation and Method: Empirical Research Methods and the Interpretive Turn,* edited by Dvora Yanow and Peregrine Schwartz-Shea, 161-82. London: Routledge.

Taylor, Charles. 1971. "Interpretation and the Sciences of Man." *The Review of Metaphysics* 25, no. 1 (September): 3-51.

Taylor, Charles. (1979) 2006. "What's Wrong with Negative Liberty." In *The Liberty Reader*, edited by David Miller 141-62. Edinburgh: Edinburgh University Press.

Yanow, Dvora. 2014. "Thinking Interpretively: Philosophical Presuppositions and the Human Sciences." In *Interpretation and Method: Empirical Research Methods and the Interpretive Turn,* edited by Dvora Yanow and Peregrine Schwartz-Shea, 5-26. London: Routledge.

Yanow, Dvora, and Peregrine Schwartz-Shea. 2014. *Interpretation and Method: Empirical Research Methods and the Interpretive Turn*. London: Routledge.

Zacka, Bernardo. 2017. *When the State Meets the Street: Public Service and Moral Agency*. Cambridge: Harvard University Press.

Zacka, Bernardo, Brooke Ackerly, Jakob Elster, Signy Gutnick Allen, Humeira Iqtidar, Matthew Longo, and Paul Sagar. 2021. "Political Theory with an Ethnographic Sensibility." *Contemporary Political Theory* 20, no. 2 (June): 385-418.

Qualitative and Multi-Method Research | Fall 2023, Volume 21.2 | https://doi.org/10.5281/zenodo.8326515

# A Sunflower Seed in a Science of Politics

Carolyn Holmes
*University of Tennessee, Konxville*

"Take these [sunflower] seeds and put them in your pockets" is a sentence that could mean several different things: offering sustenance to a hiker, preparing a gardener for a day's work, or asking someone to fill a bird feeder. Yet when spoken by a Ukrainian woman to a Russian soldier during an invasion, they became a kind of curse. "You're occupants. You're fascists. What the fuck are you doing on our land with all these guns? Take these seeds and put them in your pockets, so at least sunflowers will grow when you all lie down to die here" (Guardian 2022). A common, shared language between the woman and the soldier is what allows the confrontation to happen. But do they actually understand one another? They are in conflict, not because they cannot understand one another, but because they do. In this conversation there is a conflict in meaning, occurring within these two actors' shared language rather than a difference in language. But can a science of politics capture that conflict? Can it account for the meaning of a sunflower seed?[1]

Charles Taylor's 1971 article "Interpretation and the Science of Man" was a provocation for the entire field of political science to take seriously the significance of meaning-making activity in political life. In the fifty-plus years since the publication of this article, the study of meaning-making in political science, the meaning of "meaning," to borrow from Yanow's essay in this symposium, has become a vibrant, but still somewhat marginal, approach in political science. Taylor's article examines the idea of empiricism in a discipline that in the 1970's was rapidly experiencing a technological revolution of technology—with the advent of computer-based methods to analyze statistical inputs, in the wake of the behavioral revolution—with "the modern scientific outlook" of empiricism already having been "incorporated into the main body of the discipline" a decade before (Dahl 1961, 768, 770–71).

Taylor argues that empiricism—the collection of

1 This confrontation, a microcosm and transmutation of the invasion itself, is an illustrative example in its starkness. While certainly there are many nuances to the identities, loyalties, and interpretations of individuals within the conflict—those in multilingual households, spanning generational divides, or those with differing ties to place or power—the example is meant not as a depiction of the totality of the war or of Eastern Ukraine. There are undoubtedly complexities in the context of Ukraine and the invasion that are elided by this snapshot. Yet, it remains an important example of how the conflict played out in uncounted (uncountable?) but important ways. This uncaptured diversity may have contributed to Western Social Scientists failure to conceptualize what Ukrainian resistance would or could look like in the early days of the invasion. Instead, scholars and policy-makers predicted a short-lived and enormously asymmetrical conflict with limited local resistance, based on their "brute" data (see, e.g., Eckel 2023).

"brute data" as input, beyond subjectivity, as the result of neutral observation—was seen as a way to make scientific the study of people. Rather than having to accept first principles, or depending on an absolute inner clarity, empiricism offers another way out: (alleged) objectivity. For the empiricist, if each individual datum is not interpreted, then the collective of data is not either. As such, the application of machine-based technologies like regression analysis to large data sets "provides us with our assurance against an appeal to intuition or interpretations which cannot be understood by fully explicit procedures operating on brute data – the input" (1971, 9). This reading of the disciplinary room seems even more applicable today. Political science is increasingly computerized, with the advent of new processing capacity, greater computerized memory capacity, and new statistical software capabilities (Meyer 2022). The empiricism of the discipline is as firmly rooted, if not more so, than at the time of Taylor's writing.

I want to offer two potential extensions to Taylor's article, and his reading of the inherent tensions in an empirical science of politics. First, while Taylor pushes his readers to think about the absences in understanding created by thinking of the political world through "brute data," I will argue that even the collection of data constitutes an interpretive practice. The dataset itself constitutes an argument, in terms of what is valued and what is visible to and in later analysis. Yet this set of meaning making practices is obscured because the data themselves are not the point of inquiry, but rather the basis for analysis. Second, I will argue that there can be a useful distinction made between sharing language and sharing meaning. While Taylor focuses on the ideas of shared meaning in examining the dynamics of social scientific inquiry, I will argue that a convergence of language in practice and in analysis further obscures the differences that exist. Whether considering the technical language of the scholar or the confrontation between a woman and a soldier, being able to use the same language does not necessarily constitute a shared understanding.

## What Counts as a Seed? Data As Interpretation

In reflecting on Taylor's article in the shadow of the Russian invasion of Ukraine, I am struck by how little of the conflict is explicable through many of the expressly empirical measures we have— of diversity, of identity, of conflict. Ethnolinguistic fractionalization indices would largely fail to help us make sense of this confrontation: both the woman and the soldier spoke in Russian (Rutt 2022), and hence would not have been counted as being "observably" diverse. Those scholars who jokingly refer to themselves as the counters of guns and bombs (Youde 2019, 128) did not account for the chemistry students in Lviv making Molotov cocktails (Harding 2022), the spring mud outside of Kyiv (Tegler 2022), or the curses of Ukrainian grandmothers in the ears of Russian soldiers when evaluating the offensive and defensive military capabilities of each side.

The Russians and the Ukrainians are entering the battlefield with fundamentally different understandings of the conflict: for Putin and his acolytes, this "special military operation" is akin to a civil war, reclaiming territory from rebellious factions, territory to which Russia is entitled (RT International 2022). Putin (2022), in his speech declaring the military action, characterized the conflict by saying:

> For the United States and its allies, this is the so-called policy of containment of Russia, obvious geopolitical dividends. And for our country, this is ultimately a matter of life and death, a matter of our historical future as a people…You and I simply have not been left with any other opportunity to protect Russia, our people, except for the one that we will be forced to use today…Today's events are not connected with the desire to infringe on the interests of Ukraine and the Ukrainian people. They are connected with the protection of Russia itself from those who took Ukraine hostage.

For the Ukrainians and their allies, this war is a violation of sovereignty; an invasion of an independent country (Regan et al. 2022). In the days before the invasion, Ukrainian President Volodymyr Zelenskyy alluded to this incommensurability when he addressed the Russian people directly in a televised speech, in which he said, "Ukraine in your news and Ukraine in reality are two completely different countries. The most important difference is that ours is real" (Sonne 2022). One side frames the conflict in terms of civil war, the other in terms of interstate war. Are we, as analysts of the social world, to disregard these differences? Or can we fruitfully examine the ways in which this clash of meanings illuminates the ways in which the conflict itself is operating?

An (allegedly) empirical science of politics obscures the fact of its own interpretation and fails to recognize many meaningful dynamics in the political world. In the context of the Russian invasion of Ukraine, who was collecting and reporting supposedly brute data in the early days of the war fundamentally changed the narrative those data supposedly portrayed. Even something as apparently countable as battlefield deaths becomes a source of contention: What is a battle? Which deaths are to be counted, and by whom? Russian and Ukrainian sources reported vastly different numbers of casualties and did so, in part, to shape the discussion of the war itself (Coleman 2022). While these differences might be attributed to "spin" or "propaganda" and therefore beyond a fair critique of an empiricist position, similar discrepancies arise when considering counting such statistics in existing, and well-used Large-N datasets.

Counting apparently stable and objective phenomena, like battlefield deaths or state violence against civilians, results in enormously varied total numbers and incident counts across datasets like the Uppsala Conflict Data Program or the Correlates of War. Take, for example, the UCDP One-Sided Violence dataset (Eck and Hultman 2007), which reports civilian killings by state forces, that "records no entries for Burundi (1993), during which there were well-documented mass killings that, according to a UN Commission of Inquiry, included acts of genocide" (Broache et al. 2022).

There are many well-used datasets, from programs like Uppsala, PRIO, Systemic Peace, and others, in which the United States has a zero count of violence against civilians by state forces. Yet we know the names of Ma'Khia Bryant (Williams, Healy, and Wright 2021), Daunte Wright (New York Times 2022), Breonna Taylor (Oppel Jr., Taylor, and Bogel-Burroughs 2022), Elijah McClain (Tompkins 2022), George Floyd (Hill et al. 2020), Eric Garner (Baker, Goodman, and Mueller 2015), Michael Brown (New York Times 2014), and so many others specifically because of the violence inflicted on them by the state. Feminist scholars of international relations also challenge these purportedly objective counts of violence by noting the kinds of violence that are not counted, like intimate partner violence (Østby, Leiby, and Nordås 2019), or sexual and gender-based violence in conflict zones (True 2015). What counts as violence and what counts as death, for the purposes of these datasets, ends up making the world, as it is depicted in social scientific research.

What "counts" in the quantification of the social world—what is or is not an instance of a given phenomenon—is a meaning-making practice. Judgement calls and interpretations of the social world make certain kinds of political behaviors, whether violence or voting or diversity, visible or invisible to researchers. Far from being the objective, neutral observation that Taylor characterizes as empiricism, the data themselves constitute an implicit interpretation of the phenomena they seek to depict. This is not to say that such interpretations are dubious, or ill-intentioned. Indeed, they could be eminently defensible, but they are not beyond subjectivity.[2]

Yet the foundational assumption—that the data are objective observations, counts, or measures of fixed subjects at a point in time—persists when it comes to analyzing these data. What are scholars devoted to understanding the social world to make of the fortresses built on these foundations? What is the sense that is being made here? The messy reality of political life asserts itself to undercut the assumed order of an objective science of politics.

The danger in making such an argument is that it invites the critique the endless regression, the constant interrogation of data at the expense of more meaningful analysis. Constantly recollecting data, ever-more tailored to an individual project, raises costs and barriers to entry for researchers. But, as with most slippery-slope style arguments, such a position is clearly untenable, and logically dubious. What I propose, instead, is a re-opening of the conversation that an ambition to "brute data" analysis forecloses, to borrow from Schwartz-Shea, in this symposium. This conversation would encompass the world-making and meaning-making of data collection, and standardization, and would open up the possibilities of examining the silences and lacunae created by ambitions to data-beyond-interpretation.

## How A Sunflower Seed is Lost: Shared Language Without Shared Meaning

One of the difficulties in talking about "interpretation" in political science is the universality of the term. While those scholars explicitly involved with studying and understanding meaning making practices call themselves "interpretivists," the language of interpretation has also been used in the process of producing results and understanding the results of applying methods to apparently brute data. Empiricists in political science will often speak of having to "interpret" output from large-N analysis, regarding the fit of results with background or theoretical expectations, the marginal effects of coefficients, or similar (see, e.g., Mummolo and Peterson 2018; Keele, Stevenson, and Elwert 2020; Jordan and Philips 2023). As such, the empiricist orientation involves interpretation of statistical results, but denies the interpretation—work that goes into producing the apparently fixed data. In this mode, interpretation is a translation from one medium of communication, numbers, to another, words. The numbers themselves are not carriers of meaning, only of "significance." The sense that is made is an explanation by writer to readers of, for example, standardized coefficients, or marginal effects (King 1986).

2 In articles announcing datasets made available, the authors will often explain their methods for compiling the data, their standards for inclusion and exclusion, and their primarily sources, though not often individual judgment calls on a case or an instance (see, e.g., Birch and Muchlinski 2020; Cohen and Nordås 2014).

This process is, in some way, a sense-making, a finding of coherence, in the way that Taylor speaks of interpretation. The empiricist position starts to look and sound like the rationalist position, insofar as convincing an interlocutor that the interpretation of results is only possible if "at some point they share our understanding" (Taylor 1971, 8). In some ways, then, the chief virtue of empiricism—that it does not depend on argument or underlying assumption—has been the idea that statistical results themselves need to be interpreted. The defense offered by empiricists, that this is translation, not "interpretation," obscures what translation inevitably entails: choices that include or exclude (Evans and Fernández 2018), the use of judgment (Schedler 2012), making decisions about importance and use (Schaffner 1997), and changes in understood meaning (Schaffer 2000). This convergence of language around interpretation does not indicate a shared set of meanings or practices. Rather, it obscures the fundamental diversity in approaches to studying the social and political world.

To illustrate the point, we can circle back to the confrontation between the woman and the soldier. As Taylor (1971) points out, "a dispute [can be] at fever pitch just because both sides can fully understand the other (27–8). As with the language of interpretation shared across social science epistemologies, the shared language of the woman and the Russian solider potentially devolved into their own kind of hermeneutic circles, unintelligible to one another in meaning, but in a shared, common language.

Agents sharing the same place, whether as social science researchers or as Russian speakers in Eastern Ukraine and their direct neighbors, can or may share language without sharing meaning. When the Ukrainian woman, later in their conversation, says to the soldier, "You came to my land. Do you understand? You are occupiers," he replies, "Yes." While this may be a way to try and dismiss an angry opponent, it may also, meaningfully, be a confluence of language without a confluence of meaning. Yes, he could say, I occupy, but to liberate. Yes, I invade, to protect. It is unclear whether the shared language which makes these interactions possible is accompanied by a shared meaning of the words themselves. Indeed, in that same televised speech, Ukrainian President Volodomyr Zelenskyy alluded to this fundamental tension when he said: "Note that I am now speaking in Russian, yet no one in Russia understands what these names, streets, and events mean. This is all foreign to you. Unknown. This is our land. This is our history. What are you going to fight for? And against whom?" (Sonne 2022). Just as the interpretive scholar and the statistician share the language of interpretation, without sharing the meaning carried by those words, the woman and the soldier seem to share a language, without understanding one another.

## Finding a Sunflower Seed: Interpretation and Meaning

These two lines of inquiry—brute data as interpreted and the distinction between sharing language and sharing meaning—are extensions, rather than contradictions of Taylor's work. His fundamental concerns remain relevant and important correctives to an approach which claims to be an objective and empirical approach to the study of politics. In order to make sense of a sunflower seed, scholars of the social world need something beyond "brute data" or the unreflective interpretation of statistical outputs. Taking seriously the ways in which the phenomena we study—conflict, violence, the state, or the exercise of power—are made and made meaningful, is essential in grasping the nuance and complexity of the social world.

This assertion is not to undercut or devalue the contributions of quantitative methods or empiricism within the discipline of political science. Rather, it is an attempt to argue for the relevance and importance of studying meaning-making. Empiricists within our discipline must recognize what Taylor (1971) calls "the specificity of [their] intersubjective meanings" and their "historical specificity" to avoid treating this particular approach to the study of the political world as the only important, or indeed scientific, one (40).

So, we return to the uncounted, uncountable conflict between the woman and the soldier. As with so many other major historical junctures—like the end of the Cold War, the rise of right-wing populism, the outbreak of disease—there is an inability to capture what is happening in this conversation through brute data, nor to translate that from a measure to "results." Yet, it remains an illustrative example of the resistance of Ukraine in the face of invasion, the nature of national identity, the expression and use of power. Ultimately, we need the approach outlined by Taylor—in seriously studying the processes of meaning making and intersubjectivity—to find the meaning of a sunflower seed in a study of politics.

## References

Baker, Al, J. David Goodman, and Benjamin Mueller. 2015. "Beyond the Chokehold: The Path to Eric Garner's Death." *The New York Times*, June 13, 2015. https://www.nytimes.com/2015/06/14/nyregion/eric-garner-police-chokehold-staten-island.html.

Birch, Sarah, and David Muchlinski. 2020. "The Dataset of Countries at Risk of Electoral Violence." *Terrorism and Political Violence* 32 (2): 217–36.

Broache, Michael P., Kate Cronin-Furman, Milli Lake, and Agnes Yu. 2022. "The Uncounted Dead: Aerial Bombardment and Civilian Fatalities in Conflict Data." Paper presented at the Annual Meeting of the American Political Science Association, Montreal, Canada, September 18.

Cohen, Dara Kay, and Ragnhild Nordås. 2014. "Sexual Violence in Armed Conflict: Introducing the SVAC Dataset, 1989–2009." *Journal of Peace Research* 51 (3): .

Coleman, Julie. 2022. "Russia Says Nearly 500 of Its Troops Have Been Killed in Its Invasion, but Ukraine Says the Number Is Much Higher." *Business Insider*, March 2, 2022. https://www.businessinsider.com/russia-presents-casualties-for-war-ukraine-says-number-higher-2022-3.

Dahl, Robert A. 1961. "The Behavioral Approach in Political Science: Epitaph for a Monument to a Successful Protest." *The American Political Science Review* 55, no. 4 (December): 763–72. .

Eck, Kristine, and Lisa Hultman. 2007. "One-Sided Violence Against Civilians in War: Insights from New Fatality Data." *Journal of Peace Research* 44 (2): 233–46..

Eckel, Mike. 2023. "How Did Everybody Get The Ukraine Invasion Predictions So Wrong?" *Radio Free Europe/Radio Liberty*, February 17, 2023, sec. Ukraine. https://www.rferl.org/a/russia-ukraine-invasion-predictions-wrong-intelligence/32275740.html.

Evans, Jonathan, and Fruela Fernández. 2018. "Emancipation, Secret Histories, and the Language of Hegemony." In *The Routledge Handbook of Translation and Politics*, edited by Fruela Fernández and Jonathan Evans, 1–14. New York: Routledge.

Guardian, The. 2022. "Ukrainian Woman Offers Seeds to Russian Soldiers so 'Sunflowers Grow When They Die.'" *The Guardian,* February 25, 2022. Video, 1:10. https://www.theguardian.com/world/video/2022/feb/25/ukrainian-woman-sunflower-seeds-russian-soldiers-video.

Harding, Luke. 2022. "'I Haven't Told My Granny': Ukraine's Student Molotov Cocktail-Makers." *The Guardian*, February 28, 2022. https://www.theguardian.com/world/2022/feb/28/i-havent-told-my-granny-ukraines-student-molotov-cocktail-makers.

Hill, Evan, Ainara Tiefenthäler, Christiaan Triebert, Drew Jordan, Haley Willis, and Robin Stein. 2020. "How George Floyd Was Killed in Police Custody." *The New York Times*, June 1, 2020. https://www.nytimes.com/2020/05/31/us/george-floyd-investigation.html.

Jordan, Soren, and Andrew Q Philips. 2023. "Improving the Interpretation of Random Effects Regression Results." *Political Studies Review* 21, no. 1 (February): 210-20.

Keele, Luke, Randolph T. Stevenson, and Felix Elwert. 2020. "The Causal Interpretation of Estimated Associations in Regression Models." *Political Science Research and Methods* 8, no. 1 (January): 1–13.

King, Gary. 1986. "How Not to Lie with Statistics: Avoiding Common Mistakes in Quantitative Political Science." *American Journal of Political Science* 30, no. 3 (September): 666–87.

Meyer, Cosima. 2022. "Bringing the World to the Classroom: Teaching Statistics and Programming in a Project-Based Setting." *PS: Political Science & Politics* 55, no. 1 (January): 193–97.

Mummolo, Jonathan, and Erik Peterson. 2018. "Improving the Interpretation of Fixed Effects Regression Results." *Political Science Research and Methods* 6, no. 4 (October): 829–35.

New York Times, The. 2014. "Michael Brown's Shooting and Its Immediate Aftermath in Ferguson." *The New York Times,* August 12, 2014. https://www.nytimes.com/interactive/2014/08/12/us/13police-shooting-of-black-teenager-michael-brown.html, https://www.nytimes.com/interactive/2014/08/12/us/13police-shooting-of-black-teenager-michael-brown.html.

———. 2022. "What to Know About the Death of Daunte Wright." *The New York Times*, February 21, 2022. https://www.nytimes.com/article/daunte-wright-death-minnesota.html.

Oppel Jr, Richard A., Derrick Bryson Taylor, and Nicholas Bogel-Burroughs. 2022. "What to Know About Breonna Taylor's Death." *The New York Times*, December 12, 2022. https://www.nytimes.com/article/breonna-taylor-police.html.

Østby, Gudrun, Michele Leiby, and Ragnhild Nordås. 2019. "The Legacy of Wartime Violence on Intimate-Partner Abuse: Microlevel Evidence from Peru, 1980–2009." *International Studies Quarterly* 63, no. 1 (March): 1–14.

Putin, Vladimir. 2022. "Full Text: Putin's Declaration of War on Ukraine." *The Spectator*, February 24, 2022. https://www.spectator.co.uk/article/full-text-putin-s-declaration-of-war-on-ukraine/.

Regan, Helen, Aditi Sangal, Lianne Kolirin, and Hafsa Khalil. 2022. "Zelensky Says Ukrainian Sovereignty Extends to 'Entire Territory.'" *CNN*, June 29, 2022. https://www.cnn.com/europe/live-news/russia-ukraine-war-news-06-28-22/h_dacc4d459ece77c51a3fcf9a4074163f.

RT International. 2022. "Conflict between Russia and Ukraine a 'Civil War' of Sorts – Putin," October 27, 2022. https://www.rt.com/russia/565469-putin-ukraine-civil-war-valdai/.

Rutt, Leana. 2022. "Ethno-Linguistic Identity in Russia and Ukraine." *Confluence*, November 11, 2022. https://confluence.gallatin.nyu.edu/sections/research/ethno-linguistic-identity-in-russia-and-ukraine.

Schaffer, Frederic Charles. 2000. *Democracy in Translation: Understanding Politics in an Unfamiliar Culture*. Ithaca: Cornell University Press.

Schaffner, Christina. 1997. "Strategies of Translating Political Texts." In *Text Typology and Translation*, edited by Anna Trosborg, 119–44. Philadelphia: John Benjamins Publishing.

Schedler, Andreas. 2012. "Judgment and Measurement in Political Science." *Perspectives on Politics* 10, no. 1 (January): 21–36.

Sonne, Paul. 2022. "Ukraine's Zelensky to Russians: 'What Are You Fighting for and with Whom?'" *Washington Post*, February 24, 2022. https://www.washingtonpost.com/national-security/2022/02/23/ukraine-zelensky-russia-address/.

Taylor, Charles. 1971. "Interpretation and the Sciences of Man." *The Review of Metaphysics* 25, no. 1 (September): 3–51.

Tegler, Eric. 2022. "Have Flat Tires And Ukraine's Mud Season Stalled The Russian Column Outside Kyiv?" *Forbes*, March 6, 2022. https://www.forbes.com/sites/erictegler/2022/03/06/have-flat-tires-and-ukraines-mud-season-stalled-the-russian-convoy-outside-kyiv/.

Tompkins, Lucy. 2022. "Here's What You Need to Know About Elijah McClain's Death." *The New York Times*, January 18, 2022. https://www.nytimes.com/article/who-was-elijah-mcclain.html.

True, Jacqui. 2015. "Winning the Battle but Losing the War on Violence: A Feminist Perspective on the Declining Global Violence Thesis." *International Feminist Journal of Politics* 17 (4): 554–72.

Williams, Kevin, Jack Healy, and Will Wright. 2021. "'A Horrendous Tragedy': The Chaotic Moments Before a Police Shooting in Columbus." *The New York Times*, April 21, 2021. https://www.nytimes.com/2021/04/21/us/columbus-police-shooting-bryant.html.

Youde, Jeremy. 2019. "'Come on Down!' Pedagogical Approaches from The Price Is Right." In *Pedagogical Journeys through World Politics*, edited by Jamie Frueh, 127–36. New York: Springer.

QMMR Qualitative & Multi-Method Research

# Symposium: Emerging Methodologists Workshop

Qualitative and Multi-Method Research Fall 2023, Volume 21.2 https://doi.org/10.5281/zenodo.8418872

## Emerging Methodologists Workshop Symposium: Introduction

Diana Kapiszewski
*Georgetown University*

Hillel David Soifer
*Temple University*

The discipline of political science faces deep challenges related to diversity, equity, and inclusion. These challenges are clearly visible in the group of scholars who develop, write on, and teach research methods (Shames and Wise 2017, Barnes 2018), and qualitative and multi-method research in particular. This deficit results, in part, from a "pipeline problem" in which junior scholars from under-represented groups who are interested in and talented with such methods do not develop or write about them, and are not encouraged and actively mentored to do so. The consequent homogeneity of the group of scholars who work on these methods inhibits the emergence of new ideas and approaches and implicitly signals that those who develop and disseminate qualitative methods – i.e., those with the authority to say how research that generates and analyzes qualitative data should be conducted – must be white men.

One step toward addressing this challenge entails encouraging advanced political science graduate students and junior faculty based at U.S. institutions who are from under-represented groups to develop, publish on, and teach qualitative methods that aim at explanation, and strategies for combining qualitative and quantitative methods. The annual "Emerging Methodologists Workshop-Qualitative and Multi-Method Explanatory Research" (hereafter EMW-QMER, http://sigla.georgetown.domains/emworkshop/), supported by generous funding from the National Science Foundation's Accountable Institutions and Behavior program, seeks to contribute to that end and, more broadly, to bolster existing networks, foster new networks, and build an inclusive intellectual qualitative and multi-methods research community.

In each one-day workshop, held on the Wednesday before the annual meeting of the American Political Science Association begins, six advanced graduate students and junior faculty from under-represented groups present and receive feedback on a paper focusing on methods for collecting, generating, and analyzing qualitative data, and/or strategies for integrating qualitative and quantitative methods. Each presenter is paired with a "Methods Mentor" who works with and supports the presenter in the months preceding and following the workshop, assisting them to develop their work and move it toward peer-reviewed publication. Methods Mentors also attend the EMW-QMER.

This symposium introduces the work presented at the inaugural EMW-QMER, held in August 2023, comprising summaries of the six papers offered. This first workshop featured a fantastic set of scholars and papers examining ethics and data quality (Mitra), interviewing techniques (Morell), participant observation (Turkmen), the integration of causal effects and causal mechanisms in multi-method research (Alcocer), concept measurement (Moore), and theory reconstruction (Walton). Workshop discussion was robust, challenging, and supportive. Paper presenters are now revising their papers based on feedback from the workshop and preparing to submit them for peer review in the near future.

We are grateful to the APSA Qualitative and Multi-Method Research section for its support of this initiative, and to EMW Steering Committee members Chloe Thurston and Sheena Chestnut Greitens for their shrewd guidance and warm encouragement. We are also very thankful to the faculty who served as Methods Mentors

this year: Cathie Jo Martin, Daniel Pemstein, Benjamin Read, Ryan Saylor, Jason Seawright, and Erica Simmons.

We encourage advanced political science graduate students and junior faculty based at U.S. institutions who are writing a paper focused specifically on developing, critiquing, challenging, or enhancing a method for collecting, generating, or analyzing qualitative data, or a technique for multi-method research, to submit a proposal for the next EMW, to take place on September 4, 2024 in Philadelphia; the call for proposals will be issued in November 2023. More information on the EMW can be found here: http://sigla.georgetown.domains/emworkshop/

Barnes, Tiffany D. (2018) "Strategies for Improving Gender Diversity in the Methods Community: Insights from Political Methodologists and Social Science Research." *PS: Political Science & Politics* 51(3): 580–87.

Shames, Shauna L., and Tess Wise. (2017) "Gender, Diversity, and Methods in Political Science: A Theory of Selection and Survival Biases." *PS: Political Science & Politics* 50(3): 811–23.

Qualitative and Multi-Method Research | Fall 2023, Volume 21.2 | https://doi.org/10.5281/zenodo.8418884

# How Do Ethical Considerations Affect Data and Findings from Field Research?

Ankushi Mitra
*Georgetown University*

Field research can bring real harm to participants and communities, and a significant literature now focuses on safeguarding ethics throughout the research process (Grimm et al. 2020). However, less attention has been paid to how decisions about ethical dilemmas impact data and findings. Because all stages of research are fundamentally structured by the political contexts in which they occur, ethical considerations can affect data and findings by shaping choices about participant and question selection, documentation, and publication. Below, I briefly overview the conditions under which researchers make these decisions and discuss their potential consequences.

## Participant Selection

The dynamic and unpredictable nature of the field means that scholars may elect not to sample or interview certain individuals, households, or groups based on evolving assessments of risk. This can occur when research might bring harm to researchers and research partners, or when the research process is likely to reveal the existence, presence, or social networks of a vulnerable population to state or non-state actors (Fujii 2012). Such decisions can have important implications for data and findings. They may lead groups with specific characteristics to be systematically excluded, generate inconsistencies between what we learn about a population and what we aim to learn, or lead to conclusions that are beyond the range of the data. For example, during my fieldwork with migrants and refugees in Tunisia, black African migrants faced heightened surveillance and policing because they were perceived as disproportionately undertaking risky boat journeys to Europe. To not draw attention to them in our field sites, we interviewed people from groups less vulnerable to surveillance. However, this also meant that we likely underestimated the barriers the broader population of migrants and refugees were facing. Experiences Sub-Saharan migrants were more likely to encounter, like certain repertoires of state control or racism and xenophobia, also remained underrepresented in our data.

## Question Selection

In Tunisia, one of our goals was to understand how people made decisions about their journeys and navigated different policy regimes. Alongside us, journalists, humanitarian organizations, and security forces were also gathering data about the same population, but to different ends. Collecting certain information that may be used to surveil or coerce participants or communities is dangerous. In our case, the risk was acute for people aiming to travel to Europe, and asking about their goals and plans thus raised ethical questions. Researchers often avoid certain questions when the very act of hearing or answering them might cause psychological harm (Cronin-Furman and Lake 2018), and when answers can expose participants and communities to broader social and political risks (Wood 2006). However, omitting certain questions may lead to missing critical information and limit researchers' ability to draw conclusions and identify patterns. In our case, we learned less about some topics, like the practices of political control people attempting to reach Europe were facing, how European practices of containment operated outside European borders, and how relationships between weaker migrant-sending and powerful migrant-receiving states functioned. Missing information can also result in extrapolations and conclusions unjustified by the scope of the data. For example, because we mainly captured information about the journeys and strategies of groups that were less vulnerable to surveillance and coercion at the time, we could not be sure that our findings would apply to the broader population of migrants of interest to us.

## Documentation

Ethical concerns also inform decisions about what scholars document, and how they document information. Certain types of records, like audio or video, involve higher risks because they are more identifiable. Dilemmas arise when participants engage in sensitive behavior or hold sensitive opinions, which can bring harm if participation in research and associated data are revealed to others. Legal and political contexts further influence these decisions: researchers may refrain from documenting certain data and opt for less identifiable methods to protect against actors like the state accessing and using it (Bloemraad and Menjívar 2022). For example, four years after conducting research in Tunisia, I was working in India. Because the government had recently passed a law that led to the detention and deportation of Muslim refugees, in interviews, we did not document details about respondents' religious affiliation and community to protect Muslim interlocuters. Such decisions can make some analyses impossible, and could lead to additional issues when undocumented data do not represent the construct researchers aim to measure. For instance, we could not examine how experiences of displacement and political repression varied between identity groups; further we were unable to adequately capture whether Muslim communities and networks shape collective action and political life among refugees in distinct ways. The mode of documentation also matters. Taping interviews might lead participants to withhold information, while relying on memory or notes can lower data quality and complicating comparing responses due to different approaches and abilities to record, recall, paraphrase, or summarize information.

## Publication

These ethical dilemmas extend to publication. If respondents can identify themselves or others, data and analysis can affect psychological well-being, interpersonal or community relations, and researcher-interlocuter interactions. Other actors can access and use published information for their own purposes. Such considerations can lead researchers to withhold or delay publicizing certain results (Wood 2006). This can create "file drawer" problems by affecting the overall representativeness of findings about a topic, collective knowledge about a phenomenon or population, and allocation of resources to certain lines of inquiry. Delaying publication creates a time lag between when research is conducted and results shared, which can affect the relevance and usefulness of findings. In India, as one example, I learned about how non-governmental organizations (NGOs) assisted refugees. These data would contribute to our understanding of relationships between civil society, the state, and marginalized communities. However, I hesitated to publicize information about any organizations working with refugees. Under India's Foreign Contribution Regulation Act, NGOs must register with the state, granting authorities the power to jeopardize their legal standing for political reasons. My concern turned out to be well-founded—recently, some NGOs revealed that their ability to operate in India had been threatened by the government for aiding refugees (Sullivan and Sur 2023).

## Recommendations and Conclusions

Because decisions made at these pivotal moments can shape the trajectory of a project, scholars must evaluate ethical challenges and their social and scientific consequences throughout the research process, and use these evaluations to inform research practices and outputs.

First, it is important to delineate the boundaries of what we can and cannot know as completely as possible by placing research within the context in which decisions about data collection, analysis, and sharing were made and describing the scope, range, constraints, and limitations of the data and findings. This involves introducing ethical constraints, identifying pathways through which they affected research, and explaining their impact and specific implications for the aims, research activities, and findings of a project. This aids meaningful interpretation of the data and results, communicates research relevance, and can guide future inquiry.

Second, transparency about ethical decision-making is important. While complete disclosure is not always possible, scholars can explain how decisions were made and data analyzed (MacLean et al. 2019). This involves sharing the principles, criteria, processes, or frameworks used to identify, evaluate, and respond to ethical issues throughout the research process, and how the social and scientific consequences of these decisions were judged. This allows others to understand researchers' decision-making processes, assess the sources and impacts of potential variation in researcher choice, and share best practices.

Third, ethically important moments can become opportunities to reveal new perspectives and information, develop the next phase of research, shift the focus of a project, generate new lines of inquiry, function as metadata, and witness how power structures shape data generation. They can also push researchers to think about other ways of gathering information, like visiting other field sites, interacting with other populations, or using other methods. For these reasons, scholars should discuss productive approaches to further investigation in their research outputs. This offers pathways for advancing inquiry based on researchers' direct, relevant experience and insights.

These steps provide a starting point for further mainstreaming the process of analyzing ethical problems and parsing out and addressing their analytic implications. Such considerations are critical because, as this article demonstrates, the politics of the field can shape researcher decision-making and the data and findings from research at all stages of the process, from design to dissemination.

## References


Bloemraad, Irene, and Cecilia Menjívar. 2022. "Precarious Times, Professional Tensions: The Ethics of Migration Research and the Drive for Scientific Accountability." International Migration Review 56, no. 1: 4–32.

Cronin-Furman, Kate and Milli Lake. "Ethics Abroad: Fieldwork in Fragile and Violent Contexts." PS: Political Science & Politics 51, no. 3: 607-614.

Fujii, Lee Ann. 2012. "Research Ethics 101: Dilemmas and Responsibilities." PS: Political Science & Politics 45, no. 4: 717–23.

Grimm, Jannis J., Kevin Koehler, Ellen M. Lust, Ilyas Saliba, and Isabell Schierenbeck. 2020. Safer Field Research in the Social Sciences: A Guide to Human and Digital Security in Hostile Environments. Thousand Oaks: SAGE Publications.

MacLean, Lauren M., Elliot Posner, Susan Thomson, and Elisabeth Jean Wood. 2019. "Research Ethics and Human Subjects: A Reflexive Openness Approach." Qualitative Transparency Deliberations, Working Group Final Reports, Report I.2, APSA.

Sullivan, Daniel and Priyali Sur. 2023. "A Shadow of Refuge: Rohingya Refugees in India." Refugees International and The Azadi Project. https://www.refugeesinternational.org/reports-briefs/shadow-of-refuge-rohingya-refugees-in-india/.

Wood, Elisabeth Jean. 2006. "The Ethical Challenges of Field Research in Conflict Zones." Qualitative Sociology 29: 373–86.



Qualitative and Multi-Method Research Fall 2023, Volume 21.2 https://doi.org/10.5281/zenodo.8418896


# Balancing Standardization and Flexibility: How to Get the Most Out of Your Interviews


Sara Morell

*The College of New Jersey*


When positivist researchers use observational data, they make research design decisions that consider both standardization and accuracy. They take a theory, or a simplification of the world based on a hypothesized relationship between an independent and dependent variable, and test that theory with observational data. This requires an empirical approach that accurately represents the world, while ensuring extraneous factors don't impact the outcomes from the data. In other words, positivist researchers consider standardization, in that they want to justify that their findings are not the result of units being treated differently (King et al. 1994). Positivist researchers also consider accuracy, in that they want to justify that the data collected reflects their phenomenon of interest, in order to facilitate rich interpretation or make causal claims (Martin 2013; Mosley 2013). In my own research, I use interviews to study how the tactics candidate training organization's use impact women's political ambition. In this context, I compared organizations focused on women and organizations not focused on gender. I wanted my findings to accurately reflect the approaches used by these organizations, and I wanted to affirm that *if* responses from women's and non-gendered organizations were different, it was because of organizational approach and not differences in the interview method itself.

In interview research, accuracy is achieved through flexibility. Interview researchers may adjust their tone and question-wording to build rapport or get respondents to open up (Rubin and Rubin 1995). Researchers who carry out interviews may also take an interview in a new direction, based on something a respondent brought up (Berry 2002). The virtue of interview research comes from maximizing flexibility, because adaptable approaches can increase the accuracy of the information gained from different respondents. Interviewing lets the respondents provide as much insight and context as they would like, allowing researchers to build and test theories through the words and perspectives of the people most relevant to the question at hand. However, positivist researchers may worry about how this flexibility introduces potential threats to inference, or the possibility that your findings result from differences in the interview style across respondents and are not the result of real-world differences between the people or groups you spoke to. This introduces a tension for positivist researchers who use interview methods. How can one maximize the benefits of flexibility, for the purposes of gaining new insights and depth of information, while still minimizing the risk that are findings are being of differences in the interview approach and not differences in the people interviewed?

I propose a framework for balancing standardization and flexibility in qualitative interviewing. When researchers believe that tailoring their tactics to a particular respondent will increase their ability to generate more accurate data, then they should prioritize flexibility. When researchers believe that keeping their approach constant across units will minimize major threats to inference, then they should prioritize standardization. Overall, this approach is framed around two questions for interview researchers to consider when making decisions – **Will this decision make it more likely that the information I gather and how I interpret it accurately reflects the phenomenon of interest?** – and – **Will this decision introduce noise into my findings that could change my results, because of how that decision introduces differences in treatment across units, rather than real differences in respondents?**

Being flexible does not necessarily introduce a risk of systematic bias that correlates with outcomes of interest, and not every decision to standardize across units will prevent the researcher from gaining nuance with their insights. But the sheer number of research design choices and in-the-moment decisions inherent to interviews calls for a framework for considering the competing goals of standardization and accuracy. Interview researchers are often operating on the fly, responding in the moment to their respondents. So, a simple framework is necessary for evaluating the decisions they make in those moments, and whether those decisions are worth it for the extra accuracy potentially gained or whether they risk introducing bias to the data.

By considering whether a particular decision will improve accuracy, or whether it risks introducing bias to the outcomes, because of how differences in the interview method may drive differences in responses, researchers can decide what to do with their many small but potentially significant interview style, question wording, and ordering decisions. For example, if a researcher worries that a particular topic will prime respondents to answer future questions differently, for reasons other than the actual factors of interest, then they may want to ask that potentially priming question at the same point in the interview for all respondents. Alternatively, if a researcher learns through initial interviews about a factor that they did not anticipate that could be relevant to their question, they may want to prioritize flexibility by following their instincts and exploring this new point, even if it means diverging from their script. Similarly, researchers who use interviews may wonder whether particular changes in question wording and order that are used to build rapport and adapt the questions to a respondent's particular experiences may impact their results.

I used this framework in my own research on the role of candidate training organizations in increasing women's representation, to determine when to prioritize standardization and flexibility. For example, I reasoned that *when* in my interviews I asked questions about identity might prime different organizations differently, with non-gender organizations more likely to highlight programs that might not have been top of mind if I hadn't primed them to think about gender. So I only asked about those topics outright at the end of my interviews. But if an organization mentioned race or gender organically, I gave myself the flexibility to follow up – based on the respondent's own raising of the topic. Additionally, because I was initially unsure about which strategies candidate organizations perceive as most effective to increase candidate recruitment, I knew that questions about organizational strategy needed to be adapted to the specific organization. I was flexible with the phrasing of these questions, based on the specific strategies the organization mentioned and the particular barriers to running they referenced. I asked considerable and varied follow-ups when organizations talked about new forms of support that had not been mentioned in other interviews. This framework provided guidance for the tradeoffs built into interview design decisions, and allowed me to make defensible claims about differences between women's and non-gendered candidate organizations because I had thought through how different decision decisions improved accuracy or minimized bias.

The intentionally broad nature of this framework also allows a wide range of researchers who use interview methods to apply this approach. Some qualitative researchers know a lot about their sample beforehand and can identify major threats to inference before they start interviewing, allowing them to prioritize standardization at moments when they expect potential biases to their findings, and prioritize flexibility when they expect it will increase the depth and nuance of the information gleaned. Alternatively, researchers may learn through their initial interviews what the major threats to inference are, in which case greater *initial* flexibility allows the researcher to adapt their interview method to gather initial findings, while greater standardization later allows for confirmation of initial findings. Overall, this framework provides simple and clear questions to consider that will allow a broad range of researchers who use interviews to decide for themselves how to best prioritize flexibility and standardization within their methodology, and allow them to make defensible claims from the rich interview data they have collected.

## References


Berry, Jeffrey M. 2002. "Validity and Reliability Issues in Elite Interviewing". *PS: Political Science and Politics* 35 (4): 679–682.

King, Gary, Robert O. Keohane, Sidney Verba, and Robert O. O. Keohane. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.

Martin, Cathie Jo. 2013. "Crafting Interviews to Capture Cause and Effect". In L. Mosley (Ed.), *Interview Research in Political Science*. Cornell University Press.

Mosley, Layna. 2013. ""Just Talk to People?" Interviews in Contemporary Political Science". In L. Mosley (Ed.), *Interview Research in Political Science*, pp. 1–31. Ithaca, NY: Cornell University Press.

Rubin, Herbert J., and Irene S. Rubin. 2011. *Qualitative interviewing: The art of hearing data*. Thousand Oaks, CA: Sage Publications.



Qualitative and Multi-Method Research Fall 2023, Volume 21.2 https://doi.org/10.5281/zenodo.8418909


# Shifting Between Modes and Roles in Participant Observation


Fulya Felicity Turkmen

*University of California, Riverside*


The COVID-19 pandemic disrupted fieldwork as we knew it and forced many researchers to conduct fieldwork using digital tools, platforms, and data (see Digital Fieldwork 2021). Nevertheless, to some extent, increasing use and availability of digital fieldwork tools and platforms also "leveled the playing field," especially for younger, technologically adept, and less privileged researchers who lack funding, support systems, training, and favorable passport status that facilitate access to fieldwork Grimm (2022, 34). Since digital research practices are now here to stay, I argue that we need to go beyond considering these practices as mere ways of compensating for on-the-ground fieldwork and come up with propositions about how researchers who have limited time and resources for various reasons can integrate online and offline fieldwork in more or less structured or systematic ways. Slightly different from Murthy (2008, 839), who argues for "a balanced combination of physical and digital ethnography" while highlighting the superior nature of physical ethnography by claiming that "new media and digital forms of 'old media' are additional, valuable methods," I argue that work conducted digitally/online is not merely "additional," and both modes can be equally valuable for researchers.

Participant observation is a research methodology that might entail the active involvement of the researcher in an online or offline social, cultural, or political setting. Researchers can gain real-time insight into the context, processes, and mechanisms behind a social or political phenomenon by immersing themselves in the settings of the observed (Ross and Ross 1974, Bositis 1988, Gillespie and Michelson 2011).

In this piece, I propose ways of integrating online and offline participant observation by taking shifting modes (online and offline) and roles of the researcher into consideration. The paper is based on my experiences of studying political engagement and mobilization of emigrants from two authoritarian states, Turkey and Zimbabwe, in London, United Kingdom.

The main goal of my research is to explore how and why emigrants from authoritarian regimes politically engage with their home countries. ""What," "how," and "why" questions are central to the study of contention and that ethnographic methods are particularly well-suited to answering them" Fu and Simmons (2021, 1967).

Thus, participant observation, as a method, enables me to observe different forms of political engagement and mobilization, including demonstrations, fundraising events, petition deliveries, and elections that emigrants engage abroad to influence politics in their home countries and gain insight into their expectations and motivations to do so. Also, considering the ever-fluctuating, unpredictable, or potentially risky or dangerous nature of participant observation in illiberal or contentious political contexts, it is challenging for researchers to be on the ground for each significant event (Fu and Simmons 2021). Thus, researchers operating in such contexts can particularly benefit from integrating online and offline participant observation in their projects.

## On the ground participant observation continuum

Based on Junker (1952), Gold (1958), Uldam and McCurdy (2013), and McCurdy and Uldam (2014), participant observation can be placed on a continuum where the "complete participant" is at one end and the "complete observer" at the other. As shown in Figure 1 below, complete participant, participant-as-observer, observer-as-participant, and complete observer are all roles researchers can assume in their on-the-ground research.

In this continuum, ranging from complete participant to complete observer, the terms participant-as-observer and observer-as-participant require special attention as in-between forms of active participation. Research participants would be aware that the researcher engages with the participants to conduct her research when the researcher acts as participant-as-observer. In this role, researchers actively engage in the participants' activities, interactions, and experiences while observing and documenting what is happening. For example, the researcher might take on tasks involving helping the organization, publicizing, or mobilizing efforts (Uldam and McCurdy 2013, 945). When the researcher assumes the observer role as a participant, she primarily functions as an observer and minimizes her occasional participation in the observed setting.

However, my experiences confirm the fluidity of the field conditions during participant observation, and how the boundaries between overt/covert/insider/outsider and observation/participation can vanish based on changing and shifting dynamics, and how researchers can assume multiple roles during participant observation. Thus, researchers need to reflect on their changing roles constantly.

## Online participant observation continuum

I place online participant observation on a continuum ranging from complete participant to complete lurker, as shown in Figure 2 below.

The complete participant may use a personal account, create content, post regularly, interact with other group members, and ultimately act as an active online community member. In contrast, the complete lurker would not participate or disclose their presence yet pays attention and listens (Popovac and Fullwood 2018, Hine 2008, Adjin-Tettey et al. 2023). The researcher's lack of visibility or non-disclosure of their presence may be intentional or unintentional due to the nature of the online platform or activities. However, a significant portion of engagement with online communities, such as private groups, would need to be participatory since such communities would require the researcher to sign up, sometimes introduce themselves, and become community members (Cleland and Macleod, 2022).

There are also in-between positions of engager-as-observer and lurker-as-observer. The engager-as-observer would not primarily create content or post regularly but still engage with others by reposting, liking, and replying to what other people post. The lurker-as-observer would primarily observe others while minimally participating and disclosing their presence. Similar to on-the-ground participant observation, online participant observation also has a fluid nature (de Seta, 2020).

## Integrating In-Person and Online Participant Observation

Integrating online and in-person participant observation allows researchers to capture broader experiences and interactions and is a powerful tool for exploring how politics work in real-time. Online participant observation provides insights into virtual communities, social media interactions, and digital political participation and mobilization of physical communities (Schrooten 2012, Paechter 2013, Balsiger and Lambelet 2014, Airoldi 2018, Bluteau 2021). For example, I utilized in-person and online participant observation methods in my research to capture the multifaceted nature of emigrant-led political activities, including offline activities, such as demonstrations, and online engagement through social media platforms.

Researchers can integrate online and in-person observation differently while positioning themselves in different places on the on-the-ground and online participant observation continua. In all these options, researchers can identify the gatekeepers, those in leadership, management, and organization positions, and essential members in organizations or communities.

Following a **sequential approach**, researchers can start their data collection and planning processes with online participant observation by observing online platforms or communities formed on those platforms to gain initial insights. In this way, researchers can collect and analyze data on social networks, online interactions, discussions, and behaviors. All these data and analyses can help researchers inform their initial approach to formulating their research questions, hypotheses, and assumptions. Then, based on these initial online observations, researchers can identify and reach out to event organizers or community leaders online to explain their research and seek permission to attend events.

Following a **concurrent approach**, researchers can conduct online and in-person participant observation concurrently for observing and analyzing different aspects of a social or political phenomenon. They can select specific offline events based on online observations, such as political rallies, community meetings, or demonstrations, that align with their research goals, observe pre-event online engagement of organizers and participants, attend these events, and then return to online spaces to explore post-event online engagement.

Researchers can also combine online and in-person observations to **triangulate data or findings**. They can compare the data or findings to identify convergent and divergent patterns, commonalities, and differences or engage in member checking by returning to the online or offline communities. In this way, they can improve the validity and reliability of their findings and offer more comprehensive perspectives of the social or political phenomenon they are studying.

Finally, researchers carrying out online participant *observation* need to consider a broad range of ethical and safety-related factors not discussed in-depth in this piece, such as the appropriateness of overt or covert research in each context and platform, safety and well-being of the researcher, the blurry lines between public/private online spaces, privacy, and anonymity of users, complex dynamics and practices regarding the collection, analysis, and publication of the data based on the online platform, research topic, and other contextual factors (Berry 2004, Hine 2008, Dittrich and Kenneally 2012, Roberts 2015, Hennell et al. 2019, Winter and Lavis 2020, Di and Liu 2021, Grimm 2022, Lavorgna and Sugiura 2022).

The present paper emphasizes the need to go beyond viewing digital methods as mere substitutes for traditional fieldwork. Instead, it advocates for a structured integration of online and offline practices, specifically focusing on participant observation. The paper explores the roles within on-the-ground and online participant observation, ranging from active participant to complete observer. Furthermore, it provides insight into the dynamic and fluid nature of participant observation, emphasizing the importance of researchers continuously reflecting on their roles. Most importantly, the paper offers practical approaches—sequential, concurrent, and triangulation— for researchers seeking to merge online and offline approaches, particularly in participant observation.

## References


Adjin-Tettey, T. D., & Garman, A. 2023. "Lurking as a mode of listening in social media: motivations-based typologies." *Digital Transformation and Society,* 2(1), 11-26.

Airoldi, Massimo 2018. "Ethnography and the digital fields of social media." *International Journal of Social Research Methodology,* 21(6), 661-673.

Balsiger, P., & Lambelet, A. 2014. "Participant observation: how participant observation changes our view on social movements." In Methodological Practices in Social Movement Research, edited by Donatella Della Porta, 144-172. New York: Oxford University Press.

Berry, David M. 2004. *"Internet research: privacy, ethics and alienation: an open-source approach." Internet research*, 14(4), 323-332.

Bluteau, Joshua M. 2021. "Legitimizing digital anthropology through immersive cohabitation: Becoming an observing participant in a blended digital landscape." Ethnography, 22(2), 267-285.

Bositis, D. A. 1988. *"Some observations on the participant method." Political Behavior,* 10, 333-348.

Cleland, Jennifer, and Anna MacLeod 2022. "Disruption in the space–time continuum: Why digital ethnography matters*." Advances in Health Sciences Education,* 27(3), 877-892.

De Seta, G. 2020. "Three lies of digital ethnography." *Journal of Digital Social Research*, 2(1), 77-97.

Dittrich, David, and Erin Kenneally 2012. "The Menlo Report: Ethical principles guiding information and communication technology research." US Department of Homeland Security.

Fu, Diana, and Erica S. Simmons 2021. "Ethnographic approaches to contentious politics: The what, how, and why." *Comparative Political Studies*, 54(10), 1695-1721.

Gillespie, A., & Michelson, M. R. 2011. "Participant observation and the political scientist: Possibilities, priorities, and practicalities." *PS: Political Science & Politics,* 44(2), 261-265.

Gold, Raymond L. 1958. "Roles in sociological field observations." *Social Forces,* 36(3), 217-223.

Grimm, Jannis Julien 2022. "The Mixed Blessing of Digital Fieldwork: Digital Security and Ethical Dilemmas of Remote Research during and after the Pandemic." *Qualitative and Multi-Method Research,* 20(2), 21-25.

Hennell, Kath, Mark Limmer, and Maria Piacentini 2019. "Ethical dilemmas using social media in qualitative social research: A case study of online participant observation." *Sociological Research Online*, 25(3), 473-489.

Hine, Christine 2008. "Virtual ethnography: Modes, varieties, affordances." In *the SAGE handbook of online research methods,* edited by Nigel G. Fielding, Raymond M. Lee, and Grant Blank, 257-270. London: SAGE Publications.

Junker, B. 1952. "Some suggestions for the design of fieldwork learning experiences." In *Cases on Field Work,* edited by Everett C. Hughes. Chicago: University of Chicago Press.

Kapiszewski, Diana, Amanda Lin, Lauren MacLean, Lahra Smith, and Daniel Solomon 2021. "Digital Fieldwork." February. https://digitalfieldwork.iu.edu/.

Labaree, R. V. 2002. "The risk of 'going observationalist': negotiating the hidden dilemmas of being an insider participant observer." *Qualitative Research,* 2(1), 97-122.

Lavorgna, Anita, and Lisa Sugiura 2022. "Direct contacts with potential interviewees when carrying out online ethnography on controversial and polarized topics: A loophole in ethics guidelines." *International Journal of Social Research Methodology,* 25(2), 261-267.

Paechter, Carrie 2013. "Researching sensitive issues online: Implications of a hybrid insider/outsider position in a retrospective ethnographic study." *Qualitative Research,* 13(1), 71-86.

Popovac, Maša, and Chris Fullwood 2018. "The psychology of online lurking." In *The Oxford Handbook of Cyberpsychology,* edited by Allison Attrill-Smith, et al., 285-305. New York: Oxford University Press.

Roberts, Lynne D. 2015. "Ethical issues in conducting qualitative research in online communities." *Qualitative Research in Psychology,* 12(3), 314-325.

Ross, J. K., & Ross, M. H. 1974. "Participant observation in political research." *Political Methodology,* 1(1), 63-88.

Schrooten, Mieke 2012. "Moving ethnography online: researching Brazilian migrants' online togetherness." *Ethnic and Racial Studies,* 35(10), 1794-1809.

Uldam, J., & McCurdy, P. 2013. "Studying social movements: Challenges and opportunities for participant observation." *Sociology Compass,* 7(11), 941-951.

Wang, Di, and Sida Liu 2021. "Doing ethnography on social media: A methodological reflection on the study of online groups in China." *Qualitative Inquiry,* 27(8-9), 977-987.

Winter, Rachel, and Anna Lavis 2020. "Looking, but not listening? Theorizing the practice and ethics of online ethnography." *Journal of Empirical Research on Human Research Ethics*, 15(1-2), 55-62.

# Integrating Potential Outcomes and Causal Mechanisms to Guide Multi-Method Research

Marco Alcocer[1]
*Harvard University*

Can evidence seeking to estimate causal effects and evidence attempting to uncover causal mechanisms be integrated in multi-method research? And if so, how? While some unified frameworks have been presented to guide the integration of causal effects and causal mechanisms in a single study, these use mono-method (quantitative) frameworks and incorporate causal mechanisms that are defined as random intervening or mediating variables (Imai et al. 2011; Glynn and Quinn 2011; Humphreys and Jacobs 2015). Yet, most theory-based and qualitative scholars argue that causal mechanisms are not random variables but static, invariant factors that should be examined through within-unit qualitative methods (e.g., Hedström and Ylikoski 2010; Beach and Pedersen 2019; Goertz and Mahoney 2013; Waldner 2016). Despite this, the literature has yet to provide formal frameworks to guide multi-method research that incorporates qualitative and quantitative methods to investigate causal effects and causal mechanisms. Instead, multi-method research tends to draw on other frameworks, such as the potential outcomes (PO) and causal graphs frameworks, to informally discuss how qualitative methods can be combined with quantitative results (e.g., Psillos 2004; Paluck 2010; Seawright 2016; Goertz 2017).

This project contributes to this literature by presenting a novel unified formal framework to conduct multi-method research. In this short article, I draw on the PO framework (Neyman 1923; Rubin 1974) and extend it to incorporate invariant causal mechanisms. This framework clarifies the role of quantitative and qualitative methods when investigating causal claims in a multi-method study.

While I only present the setup of the framework here, in the larger project I use the framework to derive the role of quantitative and qualitative methods in multi-method research for some of the most popular research designs for applied researchers: including simple randomized experiments, instrumental variables, difference-in-differences, and regression discontinuity. I also discuss key implications of the framework, including the meaning of "counterfactuals" for causal mechanisms, mechanistic heterogeneity, case selection, and generalization of causal mechanisms. Potential Outcomes and Causal Mechanisms

To begin, let us discuss the PO framework by drawing from Morgan and Winship (2015). For a binary case, each unit $i$ has two potential outcome random variables, $Y_i^1$ in the treatment state and $Y_i^0$ in the control state. The individual causal effect for unit $i$ is therefore $\theta_i = Y_i^1 - Y_i^0$. For each causal state, a treatment or exposure variable $D_i$ exists, where $D_i = 1$ for units exposed to the treatment state and $D_i = 0$ for units exposed to the control state. If we assume that some mechanism exists that leads the treatment variable $D_i$ to cause a change in $Y^1$, then we have identified where causal mechanisms fit into the potential outcome framework. Accordingly, I expand the PO framework to accommodate causal mechanisms.

For unit $i$, $D_i = 1$ causes a change in $Y_i^1$ through $M_i(D_i = 1)$, which is a non-empty set of mechanisms $M_i(D_i = 1) = \{m_1, m_2, m_3, \ldots, m_n\}_i$. The set, $M_i(D_i = 1)$, has at least one mechanism, $m_n$, and if there are more than one, the mechanisms need not be mutually exclusive. In other words, $D_i = 1$ can cause $Y_i^1$ through more than one mechanism, and maybe even through *a combination* of these mechanisms. For example, $M_i(D_i = 1)$ may cause $Y_i^1$ through $m_1$, or through $m_1 \quad m_2$, or through $(m_1 \quad m_2)$ $m_3$.[2] Importantly, $D_i = 0$ does not have any mechanisms since it is not causing anything, and therefore $M_i(D_i = 0) = \varnothing$. This implies that mechanisms are only realized when $D_i = 1$.

However, the fundamental problem of causal inference is that we cannot observe both potential outcomes. That is, for each unit, only one of the potential outcomes is realized, so that the observed outcome variable is

$$Y_i = \begin{cases} Y_i^1 \text{ if } M_i(D_i = 1) \\ Y_i^0 \text{ if } M_i(D_i = 0) \end{cases} \tag{1}$$

Written differently,

$$Y_i^0 = M_i(D_i)Y_i^1 + (1 - M_i(D_i))Y_i \tag{2}$$

1 I thank Jason Seawright for his guidance, mentorship, and feedback throughout the process of this project. I thank Stephan Haggard for his feedback on early drafts of the project. I thank the participants of the 2023 Emerging Methodologists Workshop for their insights, Daniel Pemstein for his comments, and Diana Kapiszewski, Hillel David Soifer, and Qualitative and Multi-Method Research section of APSA for supporting the project.

2 Here ∧ means "and", while ∨ means "or".

Notice that because the causal states $Y_i^{\,1}$ and $Y_i^{\,0}$ are unrealized, no mechanisms exist in these states. It is only when the treatment variable $D_i$ is realized that mechanisms are also realized, but only for units where $D_i = 1$. In other words, the extended PO framework can explain how the realized potential outcome, $Y_i$, and the treatment variable, $D_i$, are random variables while mechanisms are static and invariable within each unit $i$ ($M_i$ is only realized when $D_i = 1$ and not when $D_i = 0$).

Moreover, if we believe that causal effects are probabilistic (not deterministic) or that outcomes are not monocausal, then for realized cases of $Y_i$ (where $Y_i = 1$ denotes the outcome is realized and $Y_i = 0$ means the outcome is not realized), we will likely observe four types of general observations in our realized data: $Y_i(M_i(D_i = 1)) = 1$, $Y_i(M_i(D_i = 1)) = 0$, $Y_i(M_i(D_i = 0)) = 1$, or $Y_i(M_i(D_i = 0)) = 0$. Given the framework presented here, this means that mechanisms exist for both $Y_i(M_i(D_i = 1)) = 1$ *and* $Y_i(M_i(D_i = 1)) = 0$. In the case that $Y_i(M_i(D_i = 1)) = 1$, the causal mechanisms realized by $D_i = 1$ should link $D_i$ to $Y_i$. However, in the case that $Y_i(M_i(D_i = 1)) = 0$, we should observe some factor disrupting the causal mechanisms realized by $D_i = 1$ that should have caused an effect on $Y_i$ but does not.

An important question that remains is whether the causal mechanism(s) is the same across units, or $M_i(D_i = 1) = M(D_i = 1)$ for all $i = 1, \dots, n$ (mechanistic homogeneity), whether it varies across subsets of the units (mechanistic homogeneity within subgroups and mechanistic heterogeneity across subgroups), or whether $M_i$ is unit-specific (complete mechanistic heterogeneity). This question is equivalent to asking whether there are constant causal effects ($Y_i^{\,1} - Y_i^{\,0} = \theta$ for all $i$) or heterogeneous causal effects ($Y_i^1 - Y_i^0 = \theta_i$), except focusing on mechanisms. In the social sciences our theories most often tend to assume mechanistic homogeneity or mechanistic homogeneity within a subset of units—for example, the effect of economic development on democratization varies by level of economic inequality (low, medium and high). We certainly never assume complete mechanistic heterogeneity.

While the PO framework that incorporates causal mechanisms is presented here using a binary treatment condition, $D_i$ $\{0, 1\}$, the framework can be extended to non-binary treatment. Further, for simplicity, like the basic PO setup, I also make the stable unit treatment value assumption (SUTVA).

## Future Research and Discussion

In the broader project, I take the new framework that incorporates invariant causal mechanisms and use it to derive the role of causal mechanisms when we use quantitative methods to estimate the average causal effects, including simple randomized experiments, linear regression, difference-in-differences, instrumental variables, and regression discontinuity designs. This identifies the role of causal mechanisms and qualitative methods in multi- method research when these quantitative tools are used. I also discuss key implications of the framework in detail, including "counterfactuals" for causal mechanisms, mechanistic heterogeneity, case selection criteria, and the generalizability of causal mechanisms in multi-method research.

In sum, the framework presented here provides not only a theoretical but a practical guideline for conducting multi-method research for causal claims. This framework has the potential to guide more rigorous and robust multi-method research. It also advances the ontological and epistemic underpinnings of multi-method research and contributes to the growing literature on this methodological approach.

## References


Beach, Derek and Rasmus Brun Pedersen (2019). *Process-tracing methods: Foundations and guidelines*. University of Michigan Press.

Glynn, Adam N and Kevin M Quinn (2011). Why process matters for causal inference. *Political Analysis 19* (3), 273–286.

Goertz, Gary (2017). *Multimethod research, causal mechanisms, and case studies: An integrated approach*. Princeton University Press.

Goertz, Gary and James Mahoney (2013). *A Tale of Two Cultures: Qualitative and Quantitative Research in the Social Sciences*. Princeton: Princeton University Press.

Hedström, Peter and Petri Ylikoski (2010). Causal mechanisms in the social sciences. *Annual review of sociology 36*, 49–67.

Humphreys, Macartan and Alan M Jacobs (2015). Mixing methods: A Bayesian approach. *American Political Science Review 109* (4), 653–673.

Imai, Kosuke, Luke Keele, Dustin Tingley, and Teppei Yamamoto (2011). Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies. *American Political Science Review 105* (4), 765–789.

Morgan, Stephen L and Christopher Winship (2015). *Counterfactuals and causal inference*.

Cambridge University Press.

Neyman, Jerzy (1923). On the application of probability theory to agricultural experiments. essay on principles. *Ann. Agricultural Sciences*, 1–51.

Paluck, Elizabeth Levy (2010). The promising integration of qualitative methods and field experiments. *The ANNALS of the American Academy of Political and Social Science 628* (1), 59–71.

Psillos, Stathis (2004). A glimpse of the secret connexion: Harmonizing mechanisms with counterfactuals. *Perspectives on science 12* (3), 288–319.

Rubin, Donald B (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology 66* (5), 688.

Seawright, Jason (2016). *Multi-method social science: Combining qualitative and quanti- tative tools*. Cambridge University Press.

Waldner, David (2016). Invariant causal mechanisms. *Qualitative & Multi-Method Research 14* (1-2), 28–33.

Qualitative and Multi-Method Research Fall 2023, Volume 21.2 https://doi.org/10.5281/zenodo.8418917

# Measuring Costly Concepts: Validation Samples for Measuring Many-N Cases

Sarah Moore
*Northwestern University*

Costly concepts are concepts that are expensive or otherwise resource-intensive to obtain measurement for over many cases. Costly concepts are present across the social sciences, though particularly in the subnational study of comparative politics. Subnational democracy, local-level armed group presence, and municipal corruption are all costly concepts for which measurement requires fine-grained data that may be practically impossible to collect for many units where the data are not already available to researchers. In the absence of actual measures of costly concepts, scholars will often substitute measurement by using proxy variables in empirical analyses, which causes non-random measurement error where measurements of the costly concept and proxies are not identical. This non-random measurement error means we risk conducting biased analyses when we cannot overcome the structural challenges that preclude precise measurement of costly concepts.

For example, the quantitative literature on non-state armed actors and violent conflict has overwhelmingly relied on local violence data to measure the presence of armed groups throughout a territory (for more extensive reviews of this literature, see Arjona and Castilla 2022; Vela Barón 2021)for obvious reasons, on violence. Yet, civil war is about much more than violence. We argue that the focus on violence hinders our understanding of the most common type of armed conflict in the world today. In particular, equating civil war and violence leads to (i. However, measuring armed group presence through violence fails as a proxy in ways we would easily expect given existing theory on civil war violence (Arjona 2016; Kalyvas 2000). Alternative measures of armed group presence entail gathering extensive knowledge from local experts through fieldwork (e.g., Arjona 2016; Aponte-González, Hirschel-Burns, and Uribe 2023). However, fieldwork-based approaches to measuring local-level armed group presence are incredibly expensive and thus limited to a reduced number of cases.

How do we know the extent to which a proxy can reliably substitute measures of our costly concept? How do we improve proxies or other measures when the proxy alone is unreliable? In this work, I develop methodological tools to understand the performance of existing proxies for costly concepts and inform more sophisticated measurement strategies based on the direct measurement of a subset of cases where obtainable. Here, I focus on a summary of the former, in which I develop a framework for collecting and analyzing validation samples wherein the accurate measurement of the costly concept is obtained for a set of cases to discern the performance of a proxy over three dimensions: the extent of disagreement, the variation in the disagreement, and the predictive features of the disagreement. I further assess the type of sample required to best estimate proxy performance relative to three potential options: a random sample, a stratified random sample, or a theoretically informative sample.

My overarching argument is that having at least some information about the relative performance of a potential proxy is better than uninformed analysis with said proxy. Collecting validation samples of at least a subset of cases to obtain direct measurements of a costly concept allows researchers to understand the degree to which a proxy and concept of interest converge and provides insight into the circumstances where they do not. To illustrate the proposed methodological framework and discuss the trade-offs of some of the sampling approaches available for these validation samples in the larger paper, I rely on simulated data. I use the concept of armed group presence relative to the oft-used violence proxy to motivate the data generation process for the simulation study and I present some of this illustration and my findings here.

## Proxy Performance

Three dimensions of interest characterize proxy performance relative to proxy-costly concept disagreement: extent, variation, and predictive features. *Disagreement* is measured as any case where the proxy measure and the costly concept measure are not equal. The *extent* of disagreement is the proportion of cases where there is proxy-costly concept disagreement relative to the number of measured cases. The *variation* in disagreement is the degree to which *extent* of disagreement varies across all cases and is calculated given the sample variance of the *extent* of disagreement. Lastly, *predictive features* of disagreement are potential variables that contribute to additional knowledge about the cases where there is disagreement between the proxy and the costly concept of interest. Although this could be derived several ways, an efficient way is to estimate a feature selection model to determine which of a set of specified variables are meaningful in predicting proxy-costly concept disagreement.

In measuring armed group presence, the extent of disagreement is the proportion of cases where there was violence and no presence, or where there was no violence, but armed groups were present. The variance of disagreement between violence-presence is the dispersion of cases where violence inaccurately measures presence relative to the number of cases sampled (i.e., the sample size).[1] Lastly, the predictive features of measurement disagreement between armed group presence and violence may be variables like state capacity, historical local communist organization, or economic development. In my work, I use classification trees to perform feature selection and determine which variables are important in predicting measurement disagreement. However, any appropriate modeling scheme that highlights important predictive features of disagreement is suitable.

## What type of sample is necessary?

While collecting at least some information about the performance of proxies relative to actual measures of costly concepts is helpful, this collection ought to be guided by a systematic sampling approach. So, what sampling design is best for uncovering and estimating proxy performance? I specifically test three different sampling strategies: random sampling, stratified random sampling, and theoretically informative sampling.[2] I find that in the case of the simulation study, the three different sampling strategies provide substantively similar information over the three proxy performance dimensions. Though these sampling approaches should also be tested using real world data, as I do in later work, these initial findings indicate that researchers should feel comfortable employing any systematic sampling approach among those explored here that most efficiently meets their additional data collection needs.

## Contribution and Further Work

My larger project on difficult-to-measure concepts provides scholars with a unified framework related to the challenges and existing tools for concept measurement in the social sciences and beyond, as well as where there are gaps for continued methodological improvements. In the work I summarize here, I have focused on measurement of concepts that are directly observable, though costly to measure. I have further provided a framework to assess just how bad existing measurements are and how potential case insights can inform us of the location of bad measurement. Through the larger research project, I hope to show the ways that case-based research can help to refine large-N quantitative research toward the end of expanding the utility of the multi-method toolkit.

## References


Aponte-González, Andrés F., Daniel Hirschel-Burns, and Andres Uribe. 2023. "Contestation, Governance, and the Production of Violence Against Civilians: Coercive Political Order in Rural Colombia." *Journal of Conflict Resolution*. https://doi.org/10.1177/00220027231177591

Arjona, Ana. 2016. *Rebelocracy*. New York: Cambridge University Press.

Arjona, Ana, and Juan Pablo Castilla. 2022. "The Violent Bias in the Study of Civil War." Working Paper.

Kalyvas, Stathis N. 2000. *Logic of Violence in Civil War*. Cambridge University Press.

Vela Barón, Mauricio. 2021. "Identifying Armed Group Presence Using Hidden Markov Models." Masters Thesis, Dept. of Mathematics and Statistics: Washington University in Saint Louis.


1 This is informative that *variance* dimension is not always useful and depends on the measurement properties of the underlying variables given that we are here presumably measuring violence and presence as binary variables.

2 This is a sampling approach I develop in the paper, wherein cases are stratified along primary strata of interest and then combined into secondary strata based on their theoretical likelihood. This secondary stratification helps to condense the strata allocations and eliminates the unnecessary allocation of some sampled units to primary strata combinations that are highly unlikely.

# A Unified Approach to Theory Reconstruction

Rachel Meade
Boston University
Marcus Walton
Boston University

Generalized concepts in the literature rarely align exactly with observations in the field. Yet for researchers who come across significant discrepancies between theory and practice, there is a lack of formal guidance for revising concepts. Here we draw on our own experiences, including Walton's research on social movements and Meade's research on populism. In both cases, our research emerged from the observation that a key concept, which was meant to help explain the outcome we were initially interested in, appeared to work differently in the field than what was described in the literature. Yet as emerging scholars, we lacked a methodological framework to help us center our research around this discrepancy.

Such a research design, which is centered around a type of concept 'reformation' cuts across the typical dichotomy scholars refer to as theory building versus theory testing. If theory building is research where scholars collect data to create a theory, then testing is where one uses data to determine whether a theory has explanatory power. Our approach, which we call *theory reconstruction,* focuses on a specific type of theorizing: concepts, yet differs from both theory building and theory testing. Theory reconstruction is instead about the rebuilding of existing concepts based on empirical observations, for the purposes of challenging or revising these concepts in the literature. Using examples from existing scholarship, we propose theory reconstruction as an accessible research design to highlight entrenched assumptions in the discipline and encourage more theory based research.

Outdated or unexamined assumptions constitute both a political and a methodological problem for the discipline, furthering inequities in the field while also leading to empirically deficient explanations and concepts. Whether we realize it or not, many of the categorizations and concepts used today in American political science have roots in unexamined assumptions that shape our understanding of the global south, communities of color, as well as the politics of the poor and working class.

Yet today's mainstream advice on research design discourages scholars from using their research to revise existing theories and concepts. For example, King, Keohane, and Verba (1994) caution students against revising theories on the basis of their data, warning that such adjustments should be done "rarely and with considerable discipline." (21) Moreover, Sartori (1970) famously advises against "conceptual stretching" of theories, suggesting that theories initially developed in the West should not be extrapolated beyond their original context. Yet this standard caution towards revision stands in tension with real-world practice in the discipline. For instance, as Kapiszweski et. al. (2022) argue, most political scientists who do fieldwork engage in revision based on their data throughout the research process.

Our argument builds on several recent works that highlight approaches to case selection and field work, and encourages scholars to make methodological assumptions more transparent. Scholars have described these nonstandard research paths alternatively as the "extended case method" (Burawoy 1998), "elucidating concepts" (Schaffer 2015), "casing a study" (Soss 2018; 2021), "creative comparisons" (Simmons and Smith 2021), and "iterative fieldwork" (Kapiszewski et. al. 2022), among others.

## What is Theory Reconstruction?

Here, we define theory reconstruction as a type of theory based research that uses empirical findings to challenge and revise key concepts in the literature. It is a research design with an explicit focus on rebuilding, or 'reconstructing' existing concepts using new mechanisms, categories, processes, or perspectives.

At its core, the literature that we highlight as having utilized this approach has at least two things in common. First, the authors identify a key, or "thick" (Coppedge 1999) concept, that is important to their field observations and in the relevant literature. Secondly, the authors observe that the key concept works differently in practice than how it is conceptualized in the literature. This discrepancy can be at the level of mechanisms, about the amount of variation in different instances of the concept, or about how the concept is applied. Importantly, while this approach can be used to challenge assumptions in the existing literature, it does not necessarily discredit or falsify other uses of the concept. Similarly, in terms of external validity, the researcher should be clear about the applicability of their observations to outside cases. For the sake of clarity, we have simplified theory reconstruction into three steps:

1. *Establish the discrepancy:* Using an inductive approach,

the researcher identifies a significant discrepancy between how a key concept is understood in the literature and how that concept appears in the researcher's observations, experiences, or preliminary data.

2. *Identify the Source:* Using their data, the researcher pinpoints where the conceptualization in the literature falls short and demonstrates how the concept appears to actually work in practice. Here, the researcher identifies prevailing or taken-for-granted assumptions that influence how the concept has previously been applied.

3. *Revise the Concept:* The author then develops a new conceptualization that can clarify mechanisms, provide scope conditions, or highlight the limits of the existing literature.

## Varieties of Theory Reconstruction

There are several possible ways that researchers might identify a discrepancy and attempt to reconstruct a concept based on their observations. In order to simplify these patterns we describe four different varieties of theory reconstruction: *revising*, *extending*, *narrowing*, and *disrupting*. These groupings are neither exhaustive, nor meant to be mutually exclusive categories, but simply useful distinctions between different approaches to theory reconstruction.

Revising is the broadest type of approach. Revising is when, upon close inspection, one observes that a key concept works differently in practice than how it is assumed to work in the literature. The researcher identifies the mechanisms or features that are inconsistent with the literature and develops a 'revised' concept. An example of this is James Scott's (1985) work on class relations. Using observations in a Malaysian village, Scott challenges a popular conception of class relations at the time, which assumed an ideological domination of subordinate classes. Instead, the author revises this conception of class relations, highlighting everyday forms of peasant resistance.

The other three groupings (*extending*, *narrowing*, and *disrupting*) are consistent with revising, but represent more niche approaches that are also common:

Extending is when an understudied phenomenon is found to be a good example of a key concept in the literature that it is not typically associated with. In order to address this, the researcher applies, or 'extends' the existing concept into the understudied context. One example is Soss (2018), where the author extends the concept of political participation to describe interactions between recipients and the state in the U.S. welfare system.

Narrowing is when a single, monolithic concept exhibits significant variation or contradictory features in practice. In order to address this, the researcher specifies, or 'narrows' the use of the concept, either by or dividing it into distinct subcategories, or distinguishing between the existing concept and a new one. An example of this is Soss and Weaver (2017) who 'narrow' the conception of the state into two 'faces': the first, liberal democratic face (e.g. electoral representation) and the second face of social control, noting that the second face is particularly prevalent in poor and communities of color.

Disrupting is where a common dichotomy or spectrum between different categories fails to hold up in practice. In order to address this, the researcher highlights these limitations and "disrupts" the set of existing categories, either by proposing a new category that expands the spectrum, or by demonstrating the limits of the overall concept. An example here is Linz's (1964) seminal essay on regime type, where he challenges the dichotomy between democratic and totalitarian regimes that was prevalent at the time, arguing that cases such as Franco's Spain involve aspects of both categories, but fit into neither. Instead, Linz disrupts this conceptualization of regime type, introducing the hybrid concept of an authoritarian regime.

## Conclusion

We argue for theory reconstruction as a modest, coherent framework to substantiate and encourage further explorations of theory based research. Moreover we argue for placing concepts and concept formation front and center as the premise of the analysis. Whether *revising*, *extending*, *narrowing*, or *disrupting* concepts, researchers have long used theory reconstruction to address unexamined assumptions, and open up future avenues for more in-depth analysis.

## References

Kapiszewski, Diana; Lauren M. MacLean, and Benjamin L. Read. 2022. "Dynamic Research Design: Iteration in Field-Based Inquiry." *Comparative Politics*. 54, no.4: 645.

King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.

Burawoy, Michael. 1998. "The Extended Case Method." *Sociological Theory*. 16, no.1: 4-33.

Linz, Juan. 1964. "An Authoritarian Regime: The Case of Spain." in *Cleavages, Ideologies and Party Systems*, edited by Eric Allard and Yrjo Littunen, Yrjo Helsinki: Westermark Society, 291–341.

Sartori, Giovanni. 1970. "Concept Misformation in Comparative Politics." *The American Political Science Review*, 64, no.4: 1033–1053.

Schaffer, Frederic C. 2015. *Elucidating Social Science Concepts: An Interpretivist Guide*. Routledge.

Scott James C. 1985. *Weapons of the Weak : Everyday Forms of Peasant Resistance*. New Haven: Yale University Press.

Soss, Joe. 2018. "On *Casing* a Study versus Studying a *Case*." *Qualitative and Multi-Method Research*. 16, no.1: 21-27.

Soss, Joe and Vesla Weaver. 2017. "Police Are Our Government: Politics, Political Science, and the Policing of Race-Class Subjugated Communities." *Annual Review of Political Science*. Vol. 20: 565-591

Simmons, Erica, & Nicholas Rush Smith. (Eds.). 2021. *Rethinking Comparison: Innovative Methods for Qualitative Political Inquiry*. Cambridge: Cambridge University Press.

QMMR Qualitative & Multi-Method Research

# Book Symposium: Tasha Fairfield and Andrew Charman. 2023. *Social Inquiry and Bayesian Inference*. New York: Cambridge University Press

Qualitative and Multi-Method Research | Fall 2022, Vol. 20.2 | https://doi.org/10.5281/zenodo.8326344

## A Seminal Achievement: The First Comprehensive Approach to Formal Bayesian Process Tracing

Andrew Bennett
*Georgetown University*

I am more excited about the publication of Tasha Fairfield and Andrew Charman's *Social Inquiry and Bayesian Inference: Rethinking Qualitative Research* (2022; hereafter cited in text as SIBI) than I have been about any book for many years. Even for those who prefer to use Bayesian logic informally rather than using explicit priors and likelihood ratios, SIBI greatly clarifies the Bayesian logic that underlies process tracing, and it provides clear guidance for avoiding inferential errors. As Macartan Humphreys once put it to me, Bayesian analysis makes transparent and more reliable the judgments we had to be making anyway to make causal inferences from case studies.

SIBI vaults the discussion of Bayesian process tracing forward on many fronts: how Bayesianism differs from other approaches, how to deal with complications like multiple hypotheses rather than just hypothesis H and its negation (~H or "not H"), the pros and cons of informal and formal Bayesian analysis of evidence from cases, and improvements over existing practical advice on carrying out process tracing. Above all, SIBI makes an enormous contribution by showing that Bayesian logic can in principle be used fully and transparently on every piece of evidence to adjudicate among alternative explanations of a case, even if in practice, as SIBI's authors note, it would be unwieldy to present readers with such a full and formal analysis.

Fairfield and Charman (2022) accomplish these feats while still making SIBI accessible to graduate students and useful for instructors. They provide clear guidelines, numerous exercises, and many worked examples of their approach, relegating he more technical material to appendices. As a result, SIBI is useful both for readers interested in working through all the math and those who prefer simply to understand the intuitions behind Bayesianism and follow the steps required to use its logic in process tracing, whether formally or informally.

In this brief review, I focus on SIBI's contributions on four issues that have often been misunderstood by critics and students (SIBI outlines several of these, and other common misunderstandings, 448-54). These include: 1) the distinction between the logical mutual exclusivity of hypotheses, which Bayesian inference requires, and mutual exclusivity of variables between hypotheses, which Bayesian inference does not require; 2) the number of comparisons among hypotheses vis-à-vis the evidence that are necessary as the number of hypotheses grows; 3) the relevance of evidence to *all* the alternative hypotheses rather than to just one hypothesis or another; and, 4) the value of using the log-odds form of Bayes Theorem.[1] I conclude with four issues that deserve continuing discussion and research: 1) case specific versus population level priors and scope conditions, and ways of generalizing from the process tracing results of individual case studies to wider populations; 2) advice on case selection for small-n research; 3) different ways of estimating priors and likelihood ratios; and 4) the relationship between SIBI and other approaches that use Bayesian logic (particularly Macartan Humphreys and Alan Jacobs's 2023 book, *Integrated Inferences: Causal Models for Qualitative and Mixed Method Research,* on how to use causal models and Bayesian reasoning to integrate qualitative and quantitative research).

While traditional process tracing advice has emphasized the need to consider a wide range of alternative hypotheses, SIBI puts into sharper focus the need for constructing alternative explanations for the outcome of a case that are logically mutually exclusive and also exhaustive (MEE). As Fairfield and Charman point out, if alternative explanations are not mutually exclusive, it makes little sense to ask which provides the best explanation, and it is difficult to think of how one might attach priors and likelihood ratios to overlapping hypotheses (SIBI, 86). Yet as the authors note, the requirement of mutual exclusivity has often been misunderstood as requiring that alternative hypotheses must be monocausal or include only one variable. As they note, "mutual exclusivity of hypotheses is conceptually distinct from exclusivity of their constituent independent variables, causal factors, or mechanisms" (87). Not only can hypotheses include multiple independent variables and still be logically mutually exclusive, they can also include many or all of the same independent variables and be mutually exclusive, so long as they posit different functional relationships among the variables.[2] An internal combustion engine, for example, needs oxygen, fuel, spark, and compression to function. If the engine is not working, it could be that the spark plug is somewhat fouled and the oxygen intake and fuel lines are partly clogged, or it could be a partial malfunction of both the spark plug and the fuel line, or the malfunction could be due to a number of other combinations involving one, two, three, or all four of these same features. These alternative explanations are logically mutually exclusive in that they posit different functional explanations, and they attach to different counterfactuals on what interventions would be necessary for the engine to run smoothly.

It is certainly true that it can be difficult to construct a satisfactory set of mutually exclusive hypotheses, but this is a feature of the complexity of the world and our limited understanding of it, not a consequence of using Bayesian logic. One can always make alternative explanations mutually exclusive by attaching to each of them the claim that it is the most important factor—there can only be one most important factor. It is more useful, however, to construct mutually exclusive hypotheses that have functional differences and that therefore relate more clearly to observable implications that are more likely under one hypothesis than another. One useful starting point for constructing such hypotheses is a typology of theories about causal mechanisms that I have developed. The typology includes twelve families of theories that result from the intersection of four agent-structure relations (agent→agent, agent→structure, structure→agent, and structure→structure) and three categories of explanation that are common in the social sciences (including those that focus on ideas and legitimacy, material power, and transactions costs and institutional efficiency).[3]

The challenge of constructing an exhaustive set of hypotheses, or a set whose probabilities sum to 1, is in some sense the more demanding requirement. As Fairfield and Charman point out, we can never be fully sure we have satisfied this criterion, as it is always possible that an explanation we have not thought of is the best explanation. This is why Bayesians never put 100% certainty on an explanation even if very strong evidence gets them close to 100% confidence. The most we can hope for, the authors note, is "inference to the best *existing* explanation" (SIBI, 84), but they add that we can always add new explanations and reanalyze the evidence in light of the new set of hypotheses; as they note, this is a common practice in science (85).

A second misconception that SIBI puts to rest is the idea that the number of likelihood ratios one must consider, or the comparisons one must make between hypotheses for each piece of evidence, grows combinatorially large as the number of hypotheses increases.[4] In fact, as SIBI demonstrates, it is not necessary to compare every hypothesis to every other hypothesis

1 On the first three of these issues, see Zaks 2021, the response by Bennett, Charman, and Fairfeld 2022, and the rejoinder by Zaks 2022.

2 Fairfield and Charman note a related misconception, which is the idea that alternative explanations must always or mostly make mutually exclusive predictions about evidence. In fact, alternative explanations may make observationally equivalent predictions on many pieces of evidence—they need only make different predictions in at least one actual or possible instance (SIBI, 89).

3 Bennett 2013.. Bennett and Mishkin 2023 adds to this framework theories about intra-agent mechanisms of behavior.

4 Zaks, 2021.

vis-à-vis each piece of evidence. One need only arbitrarily choose one hypothesis and compare the likelihood of evidence under that hypothesis to the likelihood of that same evidence under each of the other hypotheses, and then one has implicitly compared the likelihood of the evidence under all the hypotheses to each other. If we know the likelihood ratio of H1 to H2 and H1 to H3 for evidence E1, then we know the likelihood ratio of H2 to H3 vis-à-vis E1. The analogy I use here is that one need not weigh every item in the grocery store to know their relative weights—we can weigh how many peanuts to a watermelon and how many to a cantaloupe, and then we also know the relative weights of the cantaloupe and the watermelon without ever directly weighing one against the other. Thus, adding a new hypothesis to the existing set of hypotheses requires only one additional comparison for each piece of evidence.

A third mistake that students often make when first learning Bayesian analysis, and one that even some methodologists slip into through poor wording, is the idea that evidence is "on" or "relevant to" or "an implication of" only one hypothesis.[5] SIBI underscores that a critical feature of Bayesian inference is that evidence has some probability under every hypothesis, and it is the *relative* likelihood of the evidence under different hypotheses that determines the probative weight of the evidence.

A fourth issue that SIBI makes admirably clear, but one that nonetheless still causes some confusion among students, is the value of using the log odds form of Bayes Theorem and an associated logarithmic scale, such as the decibel (dB) scale. As Fairfield and Charman point out, using the log odds form of Bayes Theorem greatly simplifies the mathematics of summing up the inferential weights of different pieces of evidence. In addition, our sensory systems for sight, hearing, etc. follow logarithmic scales – our ears can detect small differences in loudness or air pressure between different quiet sounds, but when sounds are already loud, our ears require bigger increments of additional air compression to discern any difference. Fairfield and Charman's suggestion for using the decibel scale to assess the weight of evidence is thus eminently sensible, and they discuss at length (SIBI, 129-36) how to think about and use this scale, as well as providing a table showing equivalent dB and odds ratios (133). Even so, I have found that students require considerable practice to be able to intuitively translate among dB, odds ratios, and percentage probabilities, and practice with a more detailed conversion chart (such as this) can be helpful.

In addition to these and many other contributions, an admirable feature of SIBI is its methodological pragmatism. While I continue to encounter people who think that those of us exploring formal Bayesian process tracing are advocating excessively ambitious uses of the method, to my knowledge *literally no one* has ever advocated that formal Bayesian process tracing should be employed and written up for every piece of evidence from a case study. Fairfield and Charman are careful, both in SIBI and in their earlier work, to acknowledge the limitations of formal Bayesian process tracing and the uncertainties it entails (indeed, as they point out, Bayesian analysis can be thought of as a means to estimate the uncertainty that inevitably remains in any study, not just a method for trying to reduce it). They also point out that it would be incredibly tedious for a reader to wade through a formal analysis of every piece of evidence in a study. I expect that a range of practices is likely to emerge:

- researchers may use Bayesian insights to strengthen informal or traditional process tracing and reduce inferential errors without ever writing up a formal Bayesian analysis of evidence
- researchers might perform formal Bayesian analysis of one or a few pieces of evidence, which they may or may not present to readers in the main text, footnotes, or appendices
- researchers might do formal Bayesian analysis on much or even all of the evidence, but only present the most important parts of this analysis (the pieces of evidence with the greatest inferential weight) to readers, as well as summary conclusions of the analysis
- researchers might do full formal Bayesian analysis of all the evidence in a study, present the most important parts in a publication, and present the rest of the full formal analysis in an online appendix.

I expect the first two of these practices will be the most common, and I would be surprised if a formal Bayesian analysis of all the evidence from a case is ever published in full, even in a book-length project. Nonetheless, demonstrating that a full and formal Bayesian analysis of case study evidence is possible, as SIBI does, is tremendously important. Not only does it clarify the logic of process tracing, it also outlines that logic in a mathematical form that quantitative methodologists and researchers find compelling and legitimate.[6]

I conclude with four issues that deserve further discussion and research. First, generalizing from the results of Bayesian analysis of evidence in one case to a wider population of cases is a complex proposition. SIBI

5 On this point see Bennett, Charman, and Fairfield 2022.

6 I was once the sole qualitative methodologist at a dinner with a half-dozen faculty who teach quantitative methods. One of them asked me, perhaps just to be polite, what was new and different in qualitative methods. When I responded, "formal Bayesian process tracing," the whole group snapped to attention.

devotes chapter five to this issue and offers sensible advice, but I suspect in practice generalization is often more complicated than the examples it discusses. As Fairfield and Charman note, with considerable understatement, in social science "scope conditions are not always explicitly stated from the outset (172)." They also acknowledge that social scientists often use what they call "patchwork hypotheses," or hypotheses that "different causal logics operate in different regions of the overall scope space" (Alex George and I have called these "contingent generalizations"). My default assumption has always been that in social life there are few simple hypotheses with broad scope conditions, so "patchwork hypotheses" are the norm. SIBI outlines procedures for dealing with such hypotheses, but a further complication is that our understanding of scope conditions can change markedly during case study research because as our understanding of the mechanisms in a case change our understanding of their scope conditions often change as well. In addition, it is difficult for scholar to articulate their background knowledge of all the cases in a population, and which pieces of background knowledge they think are important will change as their understanding of mechanisms and their scope conditions change. It is still possible to parse all of this out in Bayesian terms, as SIBI does, but I expect many adjustments are necessary in applying the posteriors on hypotheses from one case study to other cases that we already know are dissimilar in many potentially important respects.

Second, SIBI has a terrific chapter on case selection in small-n research (chap. 12), providing the most comprehensive discussion I have read of all the different approaches that have been proposed. SIBI's argument is that the best criterion for case selection is expected information gain, but that we cannot assess this a priori since we don't know the evidence and likelihood ratios of a case until we gather the evidence. At the same time, the authors maintain, we can expect to learn something from almost any case. Therefore, we should not worry too much about choosing cases that have less (*a priori* unknowable) information gain than other cases we might have chosen (567), and we should be transparent and unapologetic in giving pragmatic rationales for case selection. Still, the authors provide useful Bayesian advice on case selection (567-78): diversity among cases is generally good, similarities across cases can contribute to strong tests, there is no need to avoid cases with multiple plausible causes, and model-conforming cases are good for inferences on mechanisms while deviant cases are good for building or testing higher-level theories. They also sharply critique the concept of most- and least-likely cases.

I concur with these suggestions and insights, but the argument on which I am least certain is the claim that the least/most-likely designation is entirely unworkable, and related, I have not entirely given up on the idea that we can have case-specific priors. It is possible that my somewhat different inclinations from the authors here are simply semantic. What I think of as a "case specific prior" they might call (perhaps more accurately) background information that bears on whether the case fits the scope conditions of a theory. The problem here is that I think it is difficult to articulate all of the background knowledge about both theories and cases that informs scope conditions in sufficient detail that we can treat these scope conditions as binary as SIBI suggests (584, fn 40). Indeed, the authors themselves argue that in trying to assess the scope conditions of a theory, "it makes sense to examine more cases near the boundaries of our scope space (p. 216)," which might be read as implying that our concepts of scope conditions can be probabilistic rather than binary. Or perhaps this is a mis-reading – it comes down to whether we are treating scope conditions as inherently binary, or as probabilistic in the quantum sense, and whether we are accordingly treating uncertainty mostly or only as a reflection of our incomplete understanding of scope conditions (the typical Bayesian view) or as a feature of ontologically probabilistic scope conditions.

Consider a medical example. A doctor might have pretty strong knowledge about some of the scope conditions of theories bearing on the probability that a patient who walks into their office has ovarian cancer. If they are a male, usually a piece of background information that is evident upon first sight, the probability is zero—we could pretty clearly call this a case with known or quickly updated background information that places it outside the scope conditions for any theory of ovarian cancer. But sex is not always biologically binary due to the possibility of hermaphroditism, so there is already some uncertainty for the doctor, whether we are attributing it to possible measurement error on the background conditions or uncertainty on the scope conditions (given the infrequency of hermaphroditism, for example, there may not be adequate research on the incidence of ovarian cancer for hermaphrodites). If the patient is biologically a female (again, with some uncertainty) and the doctor already knows the patient has a mutation on the BRCA1 gene, their probability of ovarian cancer is higher than that of the general population of women. But there are many other attributes of the patient on which either the research or the doctor's knowledge of theories and their scope conditions is fuzzy: age, ethnicity, general health, etc., and their incidences of ovarian cancer. Still, the doctor's general biological knowledge and the (possibly mixed) results of research might allow educated guesses on how these attributes (to some degree instantly updated on seeing the patient) might affect the patient's likelihood of having ovarian cancer, whether we are calling that a case-specific prior or a probabilistic estimate of whether the patient falls into the scope conditions of probabilistic theories about ovarian cancer. I don't think Fairfield and Charman would disagree about the logic of the inferences involved here – it may just be that people typically use the term "case specific prior" for what Fairfield and Charman I expect would call, more accurately, a combination of less-than-certain and incomplete but often quickly updated background knowledge about particular cases together with less than complete or certain knowledge about scope conditions. As my impression is that many people tend to think in terms of "case specific priors," however, it will require ongoing efforts to get them to think more precisely in the terms that SIBI uses.

Also, I would slightly qualify the authors' advice on selecting cases and writing up how we did so. They are logically correct that we need not list all known cases before choosing which ones to study, and that listing the cases *not* chosen does convey salient information on inferences from those that were studied. I would put more emphasis, however, on their pragmatic advice that it is enormously useful to list and do preliminary research on a number of salient cases (SIBI, 569). I also think that listing the cases you almost chose, but did not choose, for process tracing is useful because it can clarify the (often pragmatic reasons) for case selection and pre-empt reviewers from criticizing your case selection because you did process tracing on a particular case they think would have been fruitful.

My critiques here are modest and I agree strongly with almost everything in SIBI's discussion of case selection. Qualitative research will be much improved if researchers and reviewers come to agreement around SIBI's advice on this topic. Even so, given long-standing debates on case selection, it will take considerable discussion to get to consensus around SIBI's advice on case selection criteria, even though that advice in my view is incisive and almost entirely correct.

Third, while SIBI provides excellent advice on estimating priors and likelihood ratios,

this is a topic that deserves more research. One question that deserves experimental work, and one on which Tasha, Theo Milonopoulus, and I have made a (thus far unsuccessful) grant application, is whether crowd-sourced estimates of priors and likelihood ratios are superior to those estimated by individual scholars. This could include several variants of crowdsourcing, including experts, non-experts, individuals estimating in isolation and then aggregating their estimates, groups discussing and then estimating, etc. A key challenge here is that we don't have fully articulated, "objective," and 100% true priors and likelihood ratios against which estimates can be measured. The best approximation might be experiments with estimation by subjects from whom one extremely powerful piece of evidence about a case is withheld, but this would bear only upon whether estimates on the rest of the evidence got close to the "true" explanation, not whether estimated priors or estimated likelihood ratios on any given piece of evidence were accurate.

Finally, I would like to hear more on the authors' views, and those of Alan Jacobs and Macartan Humphreys, on the relationship between SIBI and Humphreys and Jacobs (2023)

on using causal models and Bayesian reasoning to integrate qualitative and quantitative research (Alan Jacobs's contribution to the present symposium is an excellent start on this dialogue). I don't think there are any fundamental disagreements between these books, and they are certainly not redundant. But I'd like to hear more on these authors' views, perhaps in a future symposium in this journal.

In sum, Fairfield and Charman have made an enormous contribution by outlining far more clearly than any prior work how Bayesian logic can be applied in qualitative research. SIBI is both foundational, building on a long tradition of Bayesian analysis across many fields and getting to the root of critical issues, and practical, offering clear and actionable advice for researchers. As a teacher, reviewer, and practitioner of qualitative methods, I am excited to see the ways in which it is already beginning to improve qualitative research and make it more transparent, and I hope and believe that it will have a profound effect on how qualitative research is conducted and how it is viewed by scholars working primarily with quantitative and other methods.

## References

Bennett, Andrew. 2013. The mother of all isms: Causal mechanisms and structured pluralism in International Relations theory. *European Journal of International Relations* 19 (3) 459-481.

Bennett, Andrew, Charman Andrew, and Fairfield, Tasha. 2022. Understanding Bayesianism: Fundamentals for Process Tracers. *Political Analysis, 30* (2), 298-305.

Bennett, Andrew, and Benjamin Mishkin. 2023. Nineteen kinds of theories about mechanisms that every social science graduate student should know. In Harold Kincaid and Jeroen Van Bouwel, eds., *The Oxford Handbook of Philosophy of Political Science*, 154.

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

Humphreys, Macartan, and Alan M. Jacobs. 2023. *Integrated Inferences: Causal Models for Qualitative and Mixed-Method Inference*. New York: Cambridge University Press.

Zaks, Sherry. 2021. Updating Bayesian(s): A Critical Evaluation of Bayesian Process Tracing. *Political Analysis, 29* (1), 58-74.

Zaks, Sherry. 2022. Return to the Scene of the Crime: Revisiting Process Tracing, Bayesianism, and Murder. *Political Analysis, 30* ( 2), 306-310.

Qualitative and Multi-Method Research Fall 2022, Vol. 20.2 https://doi.org/10.5281/zenodo.8326449

# Leaning In to Analytic Explicitness

Alan M. Jacobs
*University of British Columbia*

Tasha Fairfield and Andrew Charman's *Social Inquiry and Bayesian Inference* (2022) constitutes a major contribution to the advancement of qualitative methods in our discipline. The volume provides (as far as I am aware) the first extended treatment of Bayesian qualitative inference in the social sciences, covering both the conceptual underpinnings of Bayesianism and a range of issues that arise in its practical implementation. The book's guidance is elaborated with a large number of detailed applications using real data, including re-analyses of the evidence in prominent published works of qualitative political science.

As Fairfield and Charman point out, a Bayesian approach holds the promise—among other virtues—of making qualitative research considerably more analytically explicit,[1] in two related respects. First, carrying out formal Bayesian procedures allows researchers to show exactly how they have made the leap from evidence to inference. Given a stated set of priors over the hypotheses and likelihoods of the evidence under each hypothesis, it becomes straightforward for readers to see where posterior beliefs come from once the evidence is (or is not) observed. Of course, priors and likelihoods must themselves be defended, and readers might disagree about the probabilities assigned by the researcher: explicitness provides no assurance of arriving at the right or a consensual answer. But by surfacing the key premises on which inference is grounded, a formal Bayesian approach makes the analysis far more susceptible to evaluation and critique.

Second, as Fairfield and Charman also make clear, a Bayesian approach provides researchers with a principled way of *aggregating* inferences across multiple pieces of evidence. The problem of aggregating across many pieces of evidence may be modest in situations in which all or nearly all of the evidence points in the same direction. Combining observations becomes much trickier, however, when different pieces of evidence pull in different directions. How certain should we be about a hypothesis if, say, many observations line up in its favor, but a few key pieces of evidence cut against it? Conventional, informal approaches to case-study research will typically struggle with this sort of situation because they tend to lack a principled way of *weighting* observations relative to one another. If researchers are willing quantify their priors and the likelihoods of the evidence, however, formal Bayesianism offers a powerful and transparent mechanism for drawing conclusions from an arbitrarily mixed evidentiary pattern (see pp. 116-117, sec. 3.7).

Of course, no treatment of a method—even one as clear and comprehensive as this book—can fully address all problems or complications that the approach might confront. In the remainder of this essay, I will briefly raise a few issues that I think this book leaves unresolved. I will discuss, in turn, Fairfield and Charman's approach to generalizing from cases; their defense of informalism in the derivation of priors and likelihoods; and their advice on writing up formal Bayesian analyses. Particularly on the last two points, one overall theme of my comments is to suggest that Fairfield and Charman might have leaned even further than they do into Bayesianism's potential to make qualitative inference more analytically transparent and evaluable.

## How Do Inferences Travel?

Suppose I have gathered and assessed the evidence from one or a small handful of cases: what can that evidence tell me about other, perhaps similar, cases? Fairfield and Charman (2022) come closest to addressing this question in Chapter 5, where they apply their framework to the qualitative analysis of multiple cases.

1 For an excellent discussion of how different qualitative approaches, including Bayesian analysis and more informal approaches, vary in their "explicitness," see the Qualitative Transparency Deliberation working group report by Kreuzer and Parsons (2018).

In this chapter, they describe an approach in which hypotheses come with *scope conditions* attached to them. The researcher then proceeds to collect evidence from one or more cases that fall within these scope conditions. To update on the hypothesis (with its stated scope conditions), we simply add up the weight of the evidence across the cases examined, arriving at posterior odds ratios for any given pair of rival hypotheses of interest. Fairfield and Charman also consider the auxiliary problem of how to generalize *beyond* the initial scope conditions, but my concern here is with the narrower question of how we learn across cases *within* the original scope conditions.

Fairfield and Charman work through their approach, in part, with an application to Dan Slater's research on democratic mobilization in authoritarian Southeast Asia, considering three hypotheses: one focused on the role of autonomous communal elites in fostering mobilization, a second positing economic decline as the central factor, and a third centered on stolen elections. Fairfield and Charman articulate each hypothesis with the region of Southeast Asia as an explicit scope condition. They then use Slater's evidence from two cases—the Philippines and Vietnam—to update beliefs over the three hypotheses. The weight of each piece of evidence observed, regardless of the case from which it is drawn, is simply added together to yield the relevant posterior odds ratios (over any two of the three hypotheses that we might want to compare).

If I have understood the approach here correctly, because we are always updating on the hypotheses—and because these hypotheses are framed in terms of some *set* of cases, such as autocratic countries in Southeast Asia—the posterior beliefs that we generate are always understood to apply to *all* cases that fit the stated scope conditions. In Fairfield and Charman's reanalysis of Slater's data, the weight of the evidence in Vietnam and the Philippines overwhelmingly favors the communal elites hypothesis over the economic decline and stolen elections hypothesis. On my reading of Fairfield and Charman's approach to generalization, this means that we now have much greater relative confidence in the communal elites hypothesis as it applies to *all autocracies in Southeast Asia*. We should now believe communal elites to be the overwhelmingly likely cause of any mobilization that we observe in, say, autocratic Thailand or Malaysia because of the evidence observed in Vietnam and the Philippines.

It certainly seems intuitive that what we observe in Vietnam and the Philippines should affect our beliefs about other cases that share similarities to these two. But what seems odd to me is that there does not seem to be any mechanism here for distinguishing our posterior beliefs about those cases from which we *have* observed evidence from those cases from which we have *not* observed evidence. In other words, Fairfield and Charman do not appear to build in a role for uncertainty about the degree to which conclusions travel across the domain of theoretical interest. Lesson-drawing across cases seems to be automatic, the problem of generalization apparently assumed away by the declaration of a scope condition.

I am further perplexed by Fairfield and Charman's insistence that we should aggregate the weights of the evidence in exactly the same way regardless of whether the evidence all comes from within a single case or is spread across multiple cases. Either way, as long as the evidence derives from within the stated scope condition, we are simply updating on the hypothesis. Thus, for instance, there is no distinction to be made between observing, say, three highly probative, independent pieces of evidence in favor of the communal elites hypothesis (relative to its rivals) within a single case, on the one hand, and observing those same three highly probative pieces spread across three separate cases, on the other hand. I would have thought that, unless we have strong prior beliefs about the homogeneity of cases within the scope condition, we would want to shift our beliefs more strongly in favor of the communal-elites theory under the second scenario (evidence spread across cases) than under the first (evidence all within one case).

A simple thought experiment makes especially clear what Is problematic about automatic generalization across a scope-condition-defined domain. Suppose that instead of framing the Slater hypotheses as applying to autocracies in Southeast Asia, we started by framing the hypotheses as applying to *all* autocracies (and there is nothing intrinsic to the three hypotheses that makes this implausible). Despite having dramatically expanded the hypotheses' scope, there is nothing I can see in Fairfield and Charman's approach that changes how we would update on these much more general hypotheses from, say, evidence on Vietnam and the Philippines.

Defending Fairfield and Charman's approach to generalization, at least as articulated in the book, would seem to require defending very strong assumptions of exchangeability or homogeneity across cases within a given set of scope conditions. Such assumptions will not usually be tenable in social scientific applications. An alternative approach—one that would still be broadly consistent with Fairfield and Charman's framework, I think—would involve building the researcher's beliefs about heterogeneity directly into the likelihoods of the evidence, thus allowing these beliefs to condition the portability of findings across cases. Doing so would still allow for generalization and cross-case learning, but in more sensible ways. It would have us update *more* strongly about the cause of mobilization in Thailand from evidence drawn from Thailand than from evidence drawn from Vietnam. It would generate sensibly weaker generalizations across domains that the researcher believes to be highly heterogeneous than across those that are believed to be more homogeneous. And it would take into account how the evidence is distributed across the domain, including, for instance, whether we are observing similar patterns of evidence across cases that we were not *a priori* confident would exhibit similar causal relationships or mechanisms.

## Where Do Priors and Likelihoods Come From?

Fairfield and Charman's approach formalizes inference starting from the point at which the researcher states prior beliefs about the hypotheses and (relative) likelihoods of the evidence under the hypotheses. How one derives priors and likelihoods, however, is left almost completely informal. I refer the reader to Chapter 3 for Fairfield and Charman's interesting discussion of how researchers should "inhabit the world of each hypothesis" (p. 105) to informally reason their way to their likelihoods.

It is surely impossible to formalize all aspects of any research process, and I have no quibble with Fairfield and Charman's decision to limit their own formalization to the process of inference from evidence, given a set of priors and likelihoods. What I would take issue with, however, is Fairfield and Charman's defense of this choice as reflecting fundamental limits of formalization.

One way of formalizing the generation of priors and likelihoods would be to begin with a formal *theory* of the causal processes operating in the domain of interest, perhaps expressed as a probabilistic causal model (Pearl 2009). In brief, by positing prior probability distributions over exogenous conditions, one can then use the model to derive priors on the probability of alternative causal effects or processes unfolding and about the likelihood of observing a given piece of evidence under the operation of alternative effects or processes.[2]

In Chapter 9, Fairfield and Charman (2022) argue persuasively that in most social scientific contexts, as opposed to some natural-scientific domains, we are unlikely to be able to arrive at *objective* groundings of our likelihoods. In the "hard" sciences, they point out, "strong underlying theory and well-understood error models for the measurement apparatus" (441) sometimes yield unambiguous likelihood functions with strong empirical groundings. These are conditions that rarely prevail in social scientific research situations, meaning that our likelihoods will always contain a large element of subjectivity. All of this I find persuasive.

What is not obvious to me, however, is how or why the *subjectivity* of likelihoods in the social sciences speaks particularly in favor of *informalism* in the derivation of likelihoods. It is not exactly clear from the text how Fairfield and Charman see the relationship between objectivity and formalization, but they seem to elide the two concepts in arguing against formalized theories as a source of likelihoods, writing:

> We can aim to formalize theories as mathematical models in order to make them more precise, but this approach may give only a veneer of objectivity, in that the model will have to be parameterized, and then further theories and/or prior probability distributions will be needed to inform the values of those parameters, which simply pushes the subjectivity back deeper into the model. (2022, 442)

To critique the use of a model as providing "only a veneer of objectivity" is to miss a couple of the key functions of a model, even of a model built on purely subjective assumptions. For one thing, writing down a model representing the researcher's beliefs about how the world works, and from which the likelihoods are then derived, makes explicit elements of the analysis that will otherwise remain implicit. The model may represent a purely subjective set of beliefs, and thus everything that flows from the model will necessarily be model-dependent. But the formalization itself makes clear to the reader exactly what those underlying beliefs are and how they lead to the posited likelihoods—in turn, exposing those beliefs to critical evaluation. In addition, formally deriving priors and likelihoods from a single underlying model forces internal consistency among the inputs to Bayesian analysis, in a way that informal derivation is unlikely to do.

In other words, perhaps differently from Fairfield and Charman, I understand the limits to objectivity and the merits of formalization to be quite distinct issues. To my mind, building Bayesian inference atop formalized theories does not push problems deeper into the analysis; rather, it *extends* the benefits of analytic explicitness deeper into the process of scientific reasoning.

## How to Write up Qualitative Bayes?

Whatever the benefits of formalization, formalizing inference undeniably involves tradeoffs. For qualitative

2 Macartan Humphreys and I present a causal-model-based approach to Bayesian inference in a new book (Humphreys and Jacobs 2023). My point here, however, is not about the virtues of any particular approach, but about the general idea of deriving priors and likelihoods from formalized theory.

researchers, one of the steepest of these tradeoffs involves how the empirical evidence is presented.

Qualitative researchers typically deploy a narrative structure in the empirical presentation of case evidence. Narrative can provide a particularly clear way of conveying how different case observations and events are temporally and logically connected. A narrative presentation provides the reader with a contextualized, textured, and relatively holistic understanding of the case and the multiple processes unfolding within it. Moreover, a well-written narrative can be interesting and enjoyable to read.

Bayesian inference, to put it mildly, does not readily lend itself to narrative structure. In Bayesianism qualitative analysis, the holism of a case gives way to the consideration of individual pieces of evidence and their (possibly joint) likelihoods under the rival hypotheses. While it is eminently feasible in Bayesian reasoning to take account of context, temporality, and overall patterns in the evidence, narrative per se is an awkward fit with formal Bayesianism. There is thus a risk that, in adopting a Bayesian approach to qualitative inference and reaping the gains of analytical explicitness, we lose some of the benefits of more conventional modes of qualitative research presentation.

Fairfield and Charman have a proposal for squaring this circle. They recommend that authors start with the story and then go Bayesian:

> Begin with a narrative that describes, interprets, and explains the bulk of the evidence from the perspective of the hypothesis that we consider most plausible. We then proceed to consider rival hypotheses, at which point we can employ either heuristic or explicit Bayesian analysis to evaluate how strongly the evidence supports our inference....If there are some pieces of evidence that fit poorly with the narrative account (e.g., they seem fluky or inconsistent with the author's argument), these can be deferred for explicit consideration in the subsequent Bayesian hypothesis comparison. (2022, 326)

This proposal seems, on its face, to offer the best of both worlds. Those readers who prefer to consume their cases whole will get a narrative; those who prize analytical explicitness will get their priors and odds-likelihood ratios; and the Bayesian analysis is itself helpfully contextualized.

I suspect, however, that the workability of this both-and approach hinges on the researcher's uncovering a rather tidy alignment of the evidence. If the evidence largely lines up in favor of a single hypothesis—as in many of the applied illustrations in the book—then it seems quite straightforward to construct a clear narrative that "describes, interprets, and explains the bulk of the evidence from the perspective of" that hypothesis.

Yet the data are often less cooperative than that: we often end up with a collection of observations pointing in different directions. By this, I do not simply mean that we often find evidence that multiple factors helped shaped an outcome; that is a kind of complexity that can be fairly readily captured in narrative form. What I mean is that we often find a good deal of evidence *supportive* of the claim that factor $X$ mattered to an outcome, together with a good deal of evidence *undermining* the claim that $X$ mattered to the outcome. This is an evidentiary situation that is going to be a much poorer fit with narrative presentation, as there is then no dominant theoretical logic on which to lean in organizing the story. It seems a fairly tall order to construct a story of how things unfolded within a case that is clear and readable, on the one hand, but also faithful to the empirical *uncertainty* about what happened, on the other hand.

Meanwhile, as I noted at the outset, this is the kind of situation to which formal Bayesianism is ideally suited. The problem of evidentiary cacophony is a trivial one from a Bayesian perspective. When we apply the Bayesian apparatus, supporting pieces of evidence shift our beliefs in favor of a given hypothesis relative to its rivals; undermining pieces of evidence shift our beliefs away from that hypothesis; and all shifts are weighted by likelihood ratios indicating how much more or less expected the evidence is under the hypothesis than under its rivals.[3]

My concern is that an approach to the writeup that foregrounds a narrative might only be well suited to situations in which the evidence "cooperates." Or, worse, that it might tend to yield presentations that convey more confidence in the "most plausible" hypothesis than the evidence itself justifies.

To be clear, I do not have in mind a better way of squaring this presentational circle. My main point is to suggest that the trade-off between narrative structure and Bayesian logic is a steeper one than Fairfield and Charman's proposal implies; I suspect that researchers will generally have to choose which they want to prioritize. But I could well be wrong. As Fairfield and Charman's readers begin to craft their own Bayesian case studies, we will likely see much experimentation with presentational form, perhaps giving rise to inventive syntheses between narrative coherence and analytic explicitness.

3 We can, of course, also take dependencies among observations into account in the likelihood function.

**References**

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

Humphreys, Macartan, and Alan M. Jacobs. 2023. *Integrated Inferences: Causal Models for Qualitative and Mixed-Method Inference*. New York: Cambridge University Press.

Kreuzer, Marcus, and Craig Parsons. 2018. "Epistemological and Ontological Priors: Varieties of Explicitness and Research Integrity." Final Report of the Qualitative Transparency Deliberations Working Group I.1, Subgroup A. August 2018. https://static.cambridge.org/content/id/urn:cambridge.org:id:article:S1537592720001164/resource/name/S1537592720001164sup002.pdf

Pearl, Judea. 2009. *Causality: Models, Reasoning, and Inference*. 2nd edition. New York: Cambridge University Press.

Qualitative and Multi-Method Research | Fall 2022, Vol. 20.2 | https://doi.org/10.5281/zenodo.8326432

# Social Inquiry and Bayesian Inference: An "Objective" Vision for Mixed Methods Research?

Sirus Bouchat
*Northwestern University*

Social Inquiry and Bayesian Inference takes as its premise the idea that Bayesian inference has the power to redefine methodology in political science. Putting itself in the company of works like Rethinking Social Inquiry (Brady and Collier 2010), Fairfield and Charman (2022) position the book as an intervention into the reified divide between qualitative and quantitative research, seeking to elevate Bayesian inference as the unifying framework through which to reposition qualitative research on par with quantitative approaches. The specter of "subjectivity," however, haunts the project throughout, both limiting its capacity to achieve its goals of defining a unifying framework for social scientific analysis, and leaving fundamental questions about research best practices in a Bayesian approach largely unaddressed.

The comprehensive scope of Fairfield and Charman's book reflects its ambitious aim to provide a detailed accounting of how researchers should rigorously specify and evaluate social scientific hypotheses regarding (qualitative) data using Bayesian frameworks. Much of the discourse advocating for Bayesian approaches in social science remains bisected. Quantitative approaches to integrating Bayesian methods into social science research practice range from the technical (e.g., BDA3) to the informal or colloquial. Recent works like Humphreys and Jacobs (2023) increasingly leverage Bayesian reasoning to tackle ongoing challenges across quantitative and qualitative work, such as fundamental questions of causal inference.

Unlike article-length treatments, Social Inquiry and Bayesian Inference has the breadth to provide thorough descriptions of Bayesian tools and paradigms alongside illustrative examples and exercises that make it a particularly powerful teaching tool. Even so, its expansive mandate for engaging qualitative data with Bayesian methods leaves dialogue with quantitative Bayesian approaches largely implicit, or indirectly reflected in sections targeting mixed methodology. From a qualitative research perspective, this book clearly addresses a need for detailed and practical guidance on implementing research within Bayesian logics; from a quantitative perspective, this project misses opportunities for sites of linkage in part because of the conception of "mixed methods" research it invests in. Specifically, as I discuss further below, the limited discussion of prior construction and halting directives around contending with prior probabilities themselves reflects the book's staunch defense of logical and objective Bayesianism—a stance that both limits its ability to champion truly mixed methodology while also creating a perplexing tension with the goal to better integrate Bayesian methods with qualitative approaches.

## Mixed Methodology: Unified Inference?

*Social Inquiry and Bayesian Inference* traverses familiar ground in the space given to articulating differences between Bayesian and frequentist statistical paradigms, noting the fragility of frequentist approaches to any interference in research design as well as the intractability of frequentist interpretation for many questions social scientists would like to ask. By contrast, argue Fairfield and Charman, Bayesian approaches are much more flexible, enable much more nimble use of data, and allow researchers to present their findings in more digestible formats—for example, using credible intervals that have the exact interpretation (e.g., "an *x%* chance of an event occurring") that students are repeatedly cautioned against offering for traditional frequentist confidence intervals.

This fervor for Bayesian methodology in contrast to frequentist approaches lends itself, then, to contending that (logical) Bayesianism provides a unifying bridge among traditionally dichotomized qualitative and quantitative research approaches. Fairfield and Charman (2022) suggest that the unifying quality that Bayes contributes is its inferential framework, noting that the process of updating prior probabilities holds irrespective of whether our data are qualitative or quantitative (382). To contend, though, that Bayesian inference has the potential to create parity across at least the qualitative/quantitative research methods dichotomy requires that the pivotal factor creating hierarchies or status and prestige differences between qualitative and quantitative research is differences of *inference*.

To the extent that this distinction still meaningfully exists, or that an overinvestment in the idea that quantitative methods or data hold higher status remains prevalent, I would argue that charges against qualitative research coming from quantitative scholarship more often pertain to threats to inference at the level of data selection (and the informativeness or bias of those data), than any flaw in the inferential capacity of qualitative *methodology* per se. That is, a concern about a qualitative project drawing inferences based on paired case studies or ethnographic field work might lie in the case selection criteria, or in the ability of ethnographic observation and analysis to truly capture the social or political phenomena of interest to the research question. These issues have less to do with the capacity of case study analysis or ethnographic field research to draw valid inferences based on their data, but rather are concerns about the broader normative or epistemological project of research: how much should we aim for generalizability? How valid (in the substantive sense) are studies that do not identify causation?

Concerns about data quality of bias certainly *threaten* inference, but they are not critiques of inferential process. This matters for the argument Fairfield and Charman lay out for Bayesianism as a unifying paradigm of mixed methods research. While Bayesian updating itself can be leveraged in both qualitative and quantitative domains, it does not at all resolve (and perhaps in fact heightens) concerns about what qualifies as *good data*. The unifying principle of Bayesian analysis, as articulated by Fairfield and Charman—"apply Bayes' rule to update prior odds by evaluating likelihood ratios," (2022, 383)—addresses a higher order concern about having coherent reasoning practices across research designs, but when the leverage you gain to address your research question precisely comes from updating priors with respect to *data*, Bayesianism does not have any inherent capacity to resolve the qualitative vs. quantitative divide that resides primarily in concerns about the validity of the data themselves.

Fairfield and Charman come close to acknowledging this challenge later in the book, noting that "there is no clear procedure for translating complex, narrative-based, qualitative information into precise probability statements" (2022, 441–42), and subjectivity—which they seem to use interchangeably with "arbitrariness," although I disagree—likely arises as a result. Likewise, they acknowledge translating qualitative data to quantitative forms of measurement can induce noise that even careful analysis cannot undo. A truly "mixed methods" project would treat *evidence* derived qualitatively as equal with that measured and collected quantitatively, but the insistence throughout the book on objective Bayesian analysis leads directly to a maligning of subjective measurement or assessment as "arbitrary" at best (erroneous at worst). This distinction all but guarantees that qualitative scholarship remains subject to dismissal or denigration based on its *measurement* strategies or data collection enterprise, and without resolving this distinction, no amount of Bayesian inference can truly unify the epistemological divide.

## Chasing Objectivity

Throughout the book, Fairfield and Charman reinforce their allegiance to logical Bayesianism and objective Bayesianism, arguing that these paradigms are the only appropriate and consistent frameworks through which to approach data. In Chapter 9, for example, the authors reify the distinction between qualitative and quantitative research in part by appealing to disciplinary differences in the "hard" science relative to the social sciences: social sciences, they note, "study far more complex and inherently noisier systems" (2022, 441), but rather than leveraging that insight to question the fundamental construction of knowledge and knowledge-generating processes even in the more "objective" hard sciences, they reassert the need to conform as much as possible to objective aims, measurement, and likelihood specification. This defense of objectivity throughout the book, to my mind, limits the possibilities both for truly "mixed methods" research and, puzzlingly, for qualitative research in the social sciences—ostensibly at odds with the book's main goal. Nowhere is this tension, and its implications for the practical application of the approaches detailed in the book, more evident than in the discussion (or lack thereof) of specifying prior probabilities.

## Practical Advice about Priors

In contrast to the attention given to specifying and evaluating hypotheses throughout *Social Inquiry and Bayesian Inference*, specifying priors receives relatively little coverage. Priors play an interesting if vexing role in the book as a whole: while heralded as a critically important component of Bayesian analysis and championed as a distinct advantage over frequentist frameworks (e.g., regarding incorporating information from engineering reports in evaluating spacecraft reliability for Mars missions; Fairfield and Charman 2022, 377), priors are also a site of concern about undue influence, subjectivity, and bias.

The detailed guidance and options presented throughout the book for applying Bayesian frameworks (e.g., 118, table 3.1) only make sense conditional on the establishment of prior probabilities, yet *how* precisely a prospective Bayesian researcher should do this is left as an exercise to the reader. That is, although ostensibly the authors allow for priors arising from an informed position (118, table 3.1, option a), appropriately formulating such a prior is not discussed. Per objective Bayes, defining priors over rival hypotheses proceeds from a position of ignorance, and throughout the book this type of prior appears to be the favored solution (either by utilizing a variety of priors somewhat agnostically or by specifying explicitly indifferent priors). Indeed, Fairfield and Charman raise concerns that priors should not be polluted by knowledge of the research design, hypotheses, or evidence when attempting to incorporate "background information."

For both explicit Bayesian analysis and heuristic application of Bayesian logic, they argue that "carefully discussing the strengths and weaknesses of rival explanations based on existing literature" is the obvious first step in formulating and justifying prior selection (2022, 491), but nevertheless seem preoccupied with the idea that analysis might be "sloppy" or involve ad hoc speculation, and secondarily that priors arise post hoc from evidence (492). Rather than detailing procedures for systematically devising sound priors given a review of literature or extensive expertise in a subject matter domain, though, the advice hews toward equal/ignorance priors—a position meant to reflect impartiality and objectivity, but one that instead problematically reflects a direct assertion of ignorance where none truly exists. The overemphasis on avoiding biased or subjective priors further seems misplaced given the authors' acknowledgment that adjudicating multiple priors is a possible option; specifying disagreeable, unreasonable, or biased priors is not inherently problematic, so long as a clear, scientific, and transparent process for re-evaluation is possible.

Notably, Fairfield and Charman (2022) do not only prefer an objective approach on practical grounds, but rather directly position themselves in opposition to subjective Bayesianism and rigorous attempts to instantiate informative priors. They write:

> [Others] might advocate using priors that reflect the collective knowledge or current state of consensus among a relevant community of scholars. While much has been written about eliciting prior probabilities and pooling expert opinion, our logical Bayesian approach is intended to reflect the rational beliefs of the scholar conducting the research. Rather than adopting other experts' probabilities as our own, or averaging priors across multiple scholars, we should conduct our own analysis, while of course drawing on evidence supplied by previous research. In our view, consensus building can best take place subsequently, through collective debate and scrutiny of our work, whereas when assigning priors, authors can and should draw on their own specific background knowledge, which may not be shared by other scholars. (98)

Ceding this ground explicitly weakens the vision of generating a unified approach to mixed methods research, both because it undermines precisely the types of knowledge and expertise qualitative scholars are likely to have (i.e., nuanced perspectives drawn comprehensively from across sources) and because it narrows the scope of research to focus on internal consistency at expense of the broader scientific project of knowledge.

## Bayes and the Project of Scientific Knowledge

The visionary aim of *Social Inquiry and Bayesian Inference* to provide a unifying framework for social scientific research is not met with a macro perspective or broader scope for how Bayesian approaches can inform the evolution of scientific knowledge, and specifically how studies using these approaches can build on one another. The effort and attention to detailing how researchers should iterate within their own projects without compromising scientific integrity (e.g., chap. 10) is admirable, but the concern about polluting specified priors with biased information (e.g., 97–98) creates a gap in the specific guidance offered for how researchers should engage prior literature. Fairfield and Charman take for granted that researchers do literature reviews carefully (or should), and that readers will attentively correct specious priors or analyses, but absent concrete direction for incorporating previous research into prior probabilities, the book's detailed micro perspective on rigorous Bayesian inference loses its macro counterpart: a theory of knowledge-building in the social sciences.

Even with its ambitious aims for integrating social science research under a Bayesian umbrella, and its thorough exposition of how Bayesian logic can apply to qualitative data, *Social Inquiry and Bayesian Inference* leaves unaddressed how this (objective) Bayesian approach could integrate research over time, and particularly across disciplines. The authors encourage skepticism, in fact, of research that may reflect "varying degrees of subjectivity in evaluation of likelihood ratios," and directly acknowledge that this "limits what we can reasonably expect in practice" when formalizing Bayesian inferences (2022, 444). The project's dedication to "objectivity" throughout is a particular disservice to the nuance of qualitative scholarship, which does not lack in its scientific value by leveraging data or insights that defy easy quantification, but which nevertheless remains in the shadow of quantitative claims of superiority via "objectivity." Moreover, without clearly delineating how to specify pristine priors, unencumbered by external information and not overly influenced by researcher beliefs, the vision Fairfield and Charman provide for social science research remains insulated and isolated—disconnected from a narrative of how social science research can progress, and knowledge can accumulate.

## References


Brady, Henry E., and David Collier, eds. *Rethinking Social Inquiry: Diverse Tools, Shared Standards.* 2nd edition. Lanham: Rowman & Littlefield.

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

Humphreys, Macartan, and Alan M. Jacobs. 2023. *Integrated Inferences: Causal Models for Qualitative and Mixed-Method Inference.* New York: Cambridge University Press.



Qualitative and Multi-Method Research Fall 2022, Vol. 20.2 https://doi.org/10.5281/zenodo.8326462


# Bayesian Challenges to Conventional Wisdom and Practice?


Hillel David Soifer
*Temple University*


In *Social Inquiry and Bayesian Inference*, Tasha Fairfield and Andrew Charman (Fairfield and Charman 2022) seek to provide the most comprehensive foundation for qualitative research in political science by grounding it in the fundamentals of logical Bayesianism. In previously published articles (Fairfield and Charman 2017, 2019) the authors have focused on methods for identifying and evaluating evidence for within-case analysis. But the logical Bayesian approach underpins guidance for a much wider range of research tasks in both qualitative research and beyond. And it is in these areas that the book (hereafter cited in text as SIBI) is especially powerful in breaking new ground.

In this commentary, I engage with three elements of the approach in SIBI in order to think about how the book might be received and read. I begin with their overall project of developing a unifying logic of inference. Second, I reflect on how process tracing is presented in SIBI, since it is here that I expect the book will be most controversial. Third, I highlight some other, more meso-level, ways in which the book challenges the utility of existing research practices and pushes us toward what seem to me more fruitful and practical research design. In my view, these are three salutary challenges to the existing conventional wisdom in the QMMR community. Even if not all readers are persuaded by the case that Fairfield and Charman outline, there is significant value in engaging with the positions that this book outlines.

## A Unifying Logic of Inference

I want to focus first on the book's overall orientation to research. Here, Fairfield and Charman are explicit—they believe that logical Bayesianism provides a logic of inference, or more precisely <u>the</u> single logic of inference that unifies all research that seeks to advance causal implications. This is a sharp and explicit pushback against what seems to have become conventional wisdom in the QMMR community—that there are distinct logics of inference, if not even broader differences, between qualitative and quantitative research. Against the view that there are distinct logics of inference (Goertz 2017) or even distinct "cultures" underlying qualitative and quantitative research (Goertz and Mahoney 2012), Fairfield and Charman argue that the logical Bayesian framework accommodates all kinds of data, and treats it all similarly in making and evaluating inferences.

Arguably, this is the boldest and most far-reaching attempt to assert a single logic of evidence underlying all (social science) research that the political science methods community has seen in thirty years. One reading of SIBI, and I don't at all intend this to be uncharitable, is that it is a Bayesian version of KKV (King, Keohane, and Verba 1994). After all, King, Keohane, and Verba argued that there is a single logic of inference that underlies all forms of social inquiry, and that differences among types of data were no more than cosmetic. Of course, as is well known to readers of this publication, KKV was not well received among scholars oriented to qualitative research because it tried to subsume qualitative work into a broadly quantitative paradigm. We might ask, then, about the place of qualitative research in SIBI. Are qualitative scholars going to have a parallel reaction and feel taken aback because Fairfield and Charman subsume their work into a broader paradigm of inference that washes away the unique features or nature of qualitative social science?

Here, I confess that after several readings of the manuscript, I have come to sense a tension in how SIBI conceives of qualitative research. One version of the book's approach is a purely practical one: any data, whether qualitative or quantitative, single-case or cross-case, that is informative as we seek to arbitrate among hypotheses is useful, so we should be qualitative researchers when we find useful data that is qualitative. Perhaps instead of terming this view of research a practical one, we could describe it as omnivorous—SIBI argues that we should consume and integrate into our research any data—of any kind—that is useful.

But I think that at times SIBI evinces hints that the authors have commitments to particular features of a logic of inference that falls closer to the qualitative "culture" described by Goertz and Mahoney (2012). I see a commitment to qualitative research in its own right entering into the presentation through the way SIBI discuss causation itself. For example, Fairfield and Charman write that "a well-specified explanatory hypothesis should generally include some sort of causal mechanism" (SIBI, 80). This assertion is likely not controversial for the typical reader of *QMMR*, though I return below to the question of how Fairfield and Charman approach process tracing. On the other hand, this claim is certainly not fully consistent with some approaches to thinking about causation found in (certain segments of) quantitative research. In other words, SIBI seems grounded in a fundamentally qualitative tradition of how causation should be conceptualized. But this claim that good explanation "should generally include" causal mechanism is not grounded by the authors in the foundations of logical Bayesianism, and indeed it is not justified at all. And much of the book's guidance rests heavily on this claim that causal mechanisms make hypotheses better. I wonder, then, whether much of the attempt to unify qualitative and quantitative methods found in Part III of SIBI will be seen by certain communities of quantitative scholars in a way not unlike how the QMMR community saw KKV—as an attempt to assert a logic of inference that subsumed their research into a paradigm they saw as resting on foundations incompatible with their research practices. More broadly, I expect that many qualitative researchers will be pushed by SIBI to reconsider their orientation toward quantitative research, and towards the question of whether and how distinct research methods can be combined, and knowledge can cumulate across multiple, incommensurate kinds of evidence.

## Process Tracing

A related issue, of course, is how Fairfield and Charman think about within-case analysis. Here, I turn from the broad orientation of the book toward more specific research practices. While SIBI is likely to provoke strong reactions from a variety of research communities, this is an issue on which it is especially provocative. Other scholars (Bennett and Checkel 2015; Humphreys and Jacobs 2015; Mahoney 2021) have grounded within-case analysis in a framework of Bayesian updating; that is not provocative in and of itself. Nor is the application of formal Bayesian analysis in my view the novel and notable analytical move that SIBI makes. Instead, SIBI takes a clear and controversial position about what makes within-case analysis informative. Against a robust body of scholarship (Beach and Pedersen 2019) that sees within-case analysis as informative only when it traces steps in the causal process, Fairfield and Charman argue (SIBI, 405ff) that <u>any</u> information that arbitrates among hypotheses is informative. As they write: "the notion that inference entails simply tracing causal mechanisms is a narrow understanding of what constitutes evidence."

A Bayesian logic of inference, then, provides a justification for resolving a debate about the nature of process-tracing in favor of a broader and more eclectic approach to within-case analysis that is not oriented toward causal process alone. I suspect that on this issue, SIBI will face an uphill battle in persuading those committed to the alternative view to abandon their stance. But while previous scholarship that takes this more eclectic view has too often only done so implicitly rather than explicitly justifying this broader view of within-case analysis, Fairfield and Charman make the debate explicit in a salutary way.

## Mechanisms Redux

Note, however, that the position here of decentering causal mechanisms in favor of a broader-tend approach to within-case analysis is to some extent in tension with the mechanistic view of causation that (I suggested above) serves to ground the overall project of SIBI. In trying to resolve this for myself, I've come to think that rather than arguing for a mechanistic view of causation in which causal mechanisms are the *sine qua non* of making good causal claims, Fairfield & Charman may instead see causal mechanisms as one sufficient but not necessary way in which scholars can elaborate hypotheses more precisely. Since, as SIBI argues, precise and detailed hypotheses facilitate the use of evidence that arbitrates among them, causal mechanisms are one way that scholars can improve their inferences.

This view, of course, resonates quite strongly with the emphasis in KKV on maximizing the observable implications of hypotheses, acknowledging (as KKV do) that much evidence about causal mechanism is likely to be qualitative. To return to the issue raised at the start, I think there's more to be done to pin down exactly the place of the "mainstream" qualitative research worldview and its emphasis on mechanistic causation in SIBI. If the past few paragraphs here are accurate, they suggest that a certain set of qualitative scholars may see an insufficiently mechanistic view of causation in SIBI and find themselves wary of being subsumed into its unified logic of inference. Just as I argued above, however, I think that by pushing these tensions into the open, and by taking such a clear and well-grounded position on them, SIBI will push scholars to articulate their responses in ways that will move these debates forward in fruitful ways.

## Existing Research Practices

In addition to these broader issues, SIBI is likely to provoke and persuade on the more micro-level of research practices. One is the approach to case selection, discussed in the most sustained way in Chapter 12. Here, too, SIBI takes on a robust tradition in qualitative methods scholarship, arguing against many algorithmic practices of case selection in favor of a more practical set of guidelines. Second is the use of all evidence for all hypotheses. This entails among other things a move away from dismissing alternative explanations in a perfunctory fashion in a theory chapter or via claims of controlled comparison toward systematic and thoughtful engagement with alternatives.

There are of course many other points in the book that are valuable touchstones for scholars and worthy of discussion. But these two represent points on which the book raises challenges for standard practices in qualitative research and grounds those new approaches in principles of logical Bayesianism in an especially clear and sustained way. I expect that these are areas in which SIBI will influence research practice in especially far-reaching ways: I, for one, have already been advising students to take both of these practices on board in designing and carrying out their research, and I look forward to assessing the extent to which others do as well.

In closing, it should be clear that SIBI is poised to be transformative at three levels. Working backwards through this essay, we can see that it has the potential to change existing research practices, to fundamentally reshape debates about the nature of process-tracing, and to invite new conversations about whether and how social scientific inference can be unified under a single logic. That, to put it mildly, is no small accomplishment: many of us will never write anything that shapes the way so many scholars think about and carry out their work. But scholars may use this book to justify the positions they take on these three levels without fully taking on board its underlying framework of logical Bayesianism. To what extent will the authors be satisfied in moving us a bit closer to practices consistent with Bayesianism even if we don't take on board the underlying logic? Will Fairfield and Charman be content if we all act a bit more Bayesian, or is the goal here to convert us into Bayesians? I look forward to hearing their response today, and to continuing what has already been a very fruitful conversation and learning experience over the years to come.

## References


Beach, Derek, and Rasmus Brun Pedersen. 2019. *Process Tracing Methods*. Ann Arbor: University of Michigan Press.

Bennett, Andrew, and Jeffrey T. Checkel, eds. 2015. *Process Tracing: From Metaphor to Analytical Tool.* Cambridge: Cambridge University Press.

Fairfield, Tasha, and Andrew Charman. 2017. "Explicit Bayesian Analysis for Process Tracing." *Political Analysis* 25, no. 3 (July): 363–80.

———. 2019. "A Dialogue with the Data: The Bayesian Foundations of Iterative Research in Qualitative Social Science." *Perspectives on Politics* 17, no. 1 (March): 154–67.

———. 2022. *Social Inquiry and Bayesian Inference: Rethinking Qualitative Research*. Cambridge: Cambridge University Press.

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

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

Humphreys, Macartan, and Alan M. Jacobs. 2015. "Mixing Methods: A Bayesian Approach." *American Political Science Review* 109, no. 4 (November): 653–73.

King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.

Mahoney, James. 2021. *The Logic of Social Science*. Princeton: Princeton University Press.

Qualitative and Multi-Method Research | Fall 2022, Vol. 20.2 | https://doi.org/10.5281/zenodo.8326443

# Bayesian Reflections

Tasha Fairfield
*London School of Economics*

Andrew Charman
*University of California, Berkeley*

*Social Inquiry and Bayesian Inference* (Fairfield and Charman 2022) aims to share our enthusiasm for Bayesianism as a rigorous foundation for inference that can help strengthen and improve the natural intuition that qualitative scholars bring to their research. By way of introduction, Bayesian inference is a largely intuitive process that begins by assessing the *prior odds* on rival hypotheses—that is, how plausible we find one hypothesis relative to rivals—drawing on any relevant initial knowledge we possess. We proceed to gather evidence. We evaluate the inferential weight of the evidence by asking which hypothesis makes that evidence more expected, and how much more expected relative to rivals—the Bayesian term here is the *likelihood ratio* (sometimes called the Bayes factor). We then update to obtain *posterior odds* on our hypotheses—following Bayes' rule, we gain more confidence in whichever hypothesis makes the evidence more expected.[1]

We thank all the commenters for their thoughtful engagement with our ideas, many of which break with established approaches to inference in the social sciences. We are grateful for this opportunity to discuss, debate, and clarify various points.

## Narrative Analysis and Bayesian Analysis

We concur with Bennett and Jacobs that there is ample scope for experimentation in how scholars incorporate Bayesian reasoning into qualitative research. As Bennett highlights, a central premise of our book is that many benefits can accrue from learning a bit about Bayesian inference, even if readers eschew the full machinery of Bayesian probability calculus. Yet we are also more optimistic about the role of explicit Bayesian analysis than we were at the outset of the project (Fairfield and Charman 2017), in part because we have a better sense of how often inferential errors can be made in case study analysis—in particular, taking evidence that is consistent with a theory to support that theory, without asking whether the evidence might be more expected under a rival. As we move toward more consciously structuring our thinking along Bayesian principles, it makes sense to write up and present that reasoning to readers, whether as a supplement to the case narrative that it informs, or potentially even as the centerpiece of a publication.

Jacobs is right to flag the disjuncture between traditional narratives and Bayesian inference, as well as the tradeoffs that scholars may face when deciding how to bring them together— these are indeed very different ways of presenting evidence and analysis. Yet research in the discipline commonly includes distinct components that do not necessarily fit neatly together—for instance, a multi-method design might include a formal model, a frequentist statistical analysis, and a case narrative.[2] As such, we would venture that presenting a narrative account alongside an overtly Bayesian analysis should not be seen as especially unusual or unwieldy. Moreover, we can begin to bridge the gap by recognizing the specific roles that narratives and Bayesian analysis play. Narratives allow authors to use their theory to explain their cases, while Bayesian analysis serves to explicitly test the theory by assessing how well it

1 For readers who are not familiar with Bayesian analysis, this video (https://www.youtube.com/ watch?v=Qvryz4RpTX0) may provide a useful introduction.

2 As we note in Chapter 9 (Fairfield and Charman 2022), these approaches juxtapose methods that draw on incompatible epistemological foundations.

outperforms salient rivals. To the extent that we value both endeavors—using an argument to explain empirics and testing the argument against rivals—including both components has merit. Word limits obviously pose constraints for journal articles, and here authors might well need to decide which component to emphasize in the main text. But scholars who wish to foreground a narrative may still be able to include illustrative Bayesian reasoning for a few key pieces of evidence in the main text while providing a more extensive Bayesian analysis as supplemental material. Alternatively, scholars might consider publishing a traditional narrative in one venue and a fully Bayesian analysis in another venue to reach different audiences.

As for Jacobs' point that narratives work best when all or most of the evidence supports the same hypothesis over rivals, whereas Bayesian analysis is ideally suited for handling less clearcut evidence, we agree. Explicit Bayesian analysis adds the most value when the evidence is nuanced and does not all weigh in favor of the same hypothesis (Fairfield and Charman, chap. 4, 164-66). And as Jacobs notes, these are also contexts in which a narrative account may be less useful or might convey more confidence in the leading explanation than the evidence merits. One of the main advantages of Bayesian analysis in fact is to keep us from overstating our confidence, or equivalently, to make us more aware of the uncertainty that surrounds our findings. Accordingly, we fully agree that in some situations it might make sense to prioritize explicit Bayesian analysis and abandon the narrative format. Our current project on covid origins adopts precisely that approach. Here we have a case for which the evidence is remarkably and notorious mixed—some observations weigh in favor of zoonosis, some favor a lab leak, and many observations that have been salient in public debate lend, in our analysis, little if any weight to either hypothesis. It is possible to write a seemingly coherent narrative from either a zoonosis perspective or a lab leak perspective, but even presenting both narratives in tandem, as if delivering opposing arguments to a jury,[3] does little to give readers a sense of which account is more plausible, and how much more plausible given what we know so far. A fully Bayesian approach that clearly delineates and analyzes each piece of evidence in turn is far better suited for systematically aggregating the inferential contribution of multiple evidentiary observations as well as avoiding confirmation bias (e.g., forgetting to ask whether evidence that ostensibly fits with one's preferred hypothesis might be as or even more expected under the rival hypothesis).[4]

At the same time, we would like to offer a few comments on the value of conducting and presenting a Bayesian analysis even when the evidence ostensibly lines up in favor of a leading hypothesis—considerations which lead us to hope that scholars will venture beyond the two Bayes-lite approaches that Bennett flags as most likely to take hold (simply harnessing knowledge of Bayesian probability to inform intuitive analysis of evidence, or evaluating likelihood ratios for just a few key pieces of evidence). First, it can be hard to discern how decisive the evidence actually is without focusing in on specific observations and asking how expected they would be under rival hypotheses. This point goes back to the above noted risks that case narratives may convey more confidence in our conclusions than the evidence warrants. Moreover, many narratives we have read do not do a good job of distinguishing argument and inference from empirics, and the evidence presented can be too vague or overly aggregated to evaluate its inferential weight. An explicit Bayesian analysis forces us to take greater care on these fronts and may in turn help us write better narratives. Second, readers may be more skeptical of the evidence for a claim than the author, so presenting a Bayesian justification for the weight that the author attributes to the evidence may help preclude disagreements, or at least provide a framework for discussing disagreements more productively. As emphasized in Chapter 7 (Fairfield and Charman 2022), we envision that one of the most important roles for explicit Bayesian analysis lies in structuring debates about inferences and making our analysis more amenable to scrutiny (here again we agree with Jacobs).

## Process Tracing and Mechanisms

Process tracing and causal mechanisms have of course played a central role in the development of qualitative methods, and Bayesianism is often associated with process tracing in this literature. However, as Soifer highlights, our approach diverges from the notion that "tracing causal processes" or providing evidence for each step in a causal chain is adequate, or even necessary for inference to best explanation.[5] Setting out to "trace a causal process" can be an excellent way to inductively

3 See Chan and Ridley 2021, chap. 12.

4 In contexts that are not quite as ambiguous as the covid example, competing narratives, if carefully written, could prove useful for highlighting which observations fit well and which fit awkwardly with each theory, and for conveying where the respective stories seem more or less contrived or strained. But this approach would not be a substitute for systematic Bayesian analysis. Relatedly, we caution that an adversarial approach, which some have advocated, creates incentives for each side to overstate the strength of their conclusions, whereas the goal should be honest assessment of the uncertainty surrounding the conclusions.

5 We also emphasize that causal mechanisms are rarely directly observable; they are themselves a matter of inference.

devise theory.[6] But articulating a causal process inspired by the evidence we observe is not equivalent to testing our hypothesis.[7] Testing requires comparing a hypothesis to salient rivals and evaluating relative likelihoods of the evidence. This Bayesian perspective reveals that we should not limit the search for evidence to observations that bear directly on our theorized pathway from $X$ to $Y$. Instead, we should recognize that any empirical observation which is more likely under one hypothesis relative to rival(s) contributes to updating, and we should seek out any evidence for which our hypotheses make divergent predictions.[8]

We would also like to offer some clarification regarding Soifer's musings on the role of mechanisms in our work that Bayesian inference is agnostic about the meaning of causation; it is compatible with whatever philosophy one wishes to adopt. Hypotheses could be formal models; they could invoke path dependence, complex conjunctural causation, or INUS causation; they could be either deterministic or probabilistic; they could be very specific about causal processes, or they could be less detailed, depending on the research agenda and the state of knowledge in the field. All we mean when we say that hypotheses should include a "causal mechanism" is that we should aim to clarify what kind of causal story we have in mind for how, why, and when some variables $X_i$ lead to outcome $Y$.[9] That is, we should aim to give an explanation. We doubt that most scholars would disagree with that notion.[10] Even KKV (King, Keohane, and Verba 1994, 34) write that "explanation—connecting causes and effects—is the ultimate goal." Quantitative scholars may well tend to work with hypotheses that are more sparse on explanation or causal mechanisms, while qualitative scholars tend to offer more detail. And when working with nuanced and complex qualitative information from interviews, archives, or first-hand observation—which is the central concern of our book—we do indeed need to articulate hypotheses with enough specificity to be able to "mentally inhabit" the corresponding world and reason about what observations would be more expected or less expected. As such we agree with Soifer's interpretation that expounding causal processes or mechanisms serves to make our hypotheses more precise. It is worth emphasizing that specifying hypotheses can be an iterative process; we may start a research project with rather bare-boned hypotheses and revise them to provide more detailed explanations or causal pathways as we learn more. Our Bayesian approach is accordingly compatible with research that begins by "soaking and poking," with only tentative initial ideas about possible explanations.

## Cross-Case Analysis

Although Bayesianism has most often been associated with process tracing and within-case analysis, we argue that in a Bayesian framework, there are no fundamental distinctions between within-case analysis and cross-case analysis. Whether we are studying a single case or multiple cases, all evidence contributes to inference in the same manner—by evaluating likelihoods under rival hypotheses. To recapitulate our approach, a well-articulated hypothesis should include a statement of its scope, beyond which it makes no predictions. Observations from any case that falls within the stated scope of the hypotheses under comparison then contribute some weight of evidence to the inference. In the same way that inferential weight accumulates for each evidentiary observation pertaining to a single case, the inferential weights of multiple pieces of evidence aggregate across cases and contribute additively to the posterior log-odds on the hypotheses.[11] Inferences are always provisional and comparative, in the sense that (i) posterior odds reflecting what we have learned from cases already examined become "prior odds" when moving forward to consider new cases, but what we discover in new cases may well change our view about the relative plausibility of alternative explanations, and (ii) we are always free to devise new or refined hypotheses to compare.

Four points may help to clarify our approach with respect to ~~regarding~~ Bennett's and Jacobs' queries about learning across cases. First, hypotheses, including their scope conditions, must be propositions with well-defined, if imperfectly known, truth values. A scope

6 Tracing a causal process may also be an effective way to deploy theory to explain a case.

7 In our view, some of the literature on process tracing and causal mechanisms conflates hypothesis generation with hypothesis testing (see *Qualitative & Multi-Method Research* 18(2)), while qualitative research that invokes process tracing as its methodological foundation often engages less in theory testing than in proposing a theory and using it to explain a case.

8 Nor do we necessarily need to examine evidence pertaining to every granular step in the causal chain, particularly if the hypotheses under consideration do not make strongly divergent predictions at some steps.

9 There is of course an ample literature that debates what exactly causal mechanisms are and what their relation is to inference (e.g., *Qualitative & Multi-Method Research* 14(1)), which we regard as largely beside the point from a Bayesian perspective.

10 Some might however take issue with our use of the term "causal mechanism," which is sometimes associated with deterministic causation, whereas we expect that probabilistic models of causation are more realistic and useful for most social science contexts.

11 Here we are invoking the log-odds version of Bayes' rule: the posterior log-odds equal the prior log-odds plus the net weight of evidence (Fairfield and Charman 2002, chap. 4).

condition itself is a logical proposition with some binary truth value that defines the contexts in which the hypothesis makes predictions, versus contexts in which the hypothesis makes no predictions at all. We may have epistemic uncertainty as to whether a case satisfies the stated scope, in that we do not have enough information about the case to know for sure.[12] But we do not take scope conditions to involve any intrinsic, aleatoric uncertainty.[13] Any uncertainty about "the degree to which conclusions travel across the domain of theoretical interest" (Jacobs, this symposium) is reflected in the probabilities of the articulated hypotheses with their stated scope conditions—which are part and parcel of the hypotheses themselves—just as these probabilities reflect uncertainty about any other aspects of the hypotheses, namely, the causal logics or mechanisms they propose.

Second, while hypotheses do need to contain clearly articulated scope conditions before applying the Bayesian inferential apparatus, scope does not need to be rigidly determined at the outset of research. As we learn more, we can always revise the scope conditions in our hypotheses to either pose them at higher levels of generality or restrict their predictions to narrower contexts, in accord with Bennett's observation that our understanding of scope can change substantially over the course of research. We view the complications Bennett emphasizes on this front as part of the usual give-and-take of iterative theory building and testing. Analyses can be revisited, observations can be analyzed differently or more deeply, different parts of our background information may become more or less relevant, and both theorized scope conditions and causal mechanisms can be tweaked.[14]

Third, priors and posteriors are associated with the hypotheses under comparison and necessarily match the stated scope conditions that the hypotheses articulate. Regarding Bennett's query about case-specific priors and Jacobs' question about distinguishing posterior beliefs about cases for which we have observed evidence from posteriors about cases from which we have not yet observed evidence, our response is that case-specific hypotheses have case-specific priors and case-specific posteriors; whereas hypotheses with broader scope have priors that are informed by all salient background knowledge possessed about each of the cases within its scope, and posteriors that draw on all evidence learned from any case within the scope. Medical examples like the one Bennett introduces are best understood as using (rather than testing) theories to diagnose or make prognostic predictions for an individual case (a patient),[15] as are examples of generating and assessing hypotheses about an individual case (e.g., the patient has ovarian cancer vs. irritable bowel syndrome). As for our social science example on democratic mobilization, when comparing hypotheses that are scoped to make predictions throughout Southeast Asia, logically speaking we cannot ask about priors or posteriors that apply only to some subset of Southeast Asian countries vs. priors or posteriors that apply to some other subset thereof. That is, different cases that fall within the scope of the hypotheses under comparison cannot have different priors or posteriors.[16]

Fourth, a hypothesis that makes predictions within a given theoretical domain or scope need not assert causal homogeneity across the entirely of that domain. A hypothesis can apply one causal logic within some subregion of its scope space and another distinct causal logic within some other subregion of its scope. "Patchwork" hypotheses of this sort assert causal heterogeneity, while still making predictions across all cases within their scope. (While some readers may tend to associate scope with a particular causal mechanism or casual logic, we emphasize again that the scope of a *hypothesis* is simply a statement about the contexts in which it makes predictions of any kind, vs. those in which it makes no predictions.) As we study more cases or expand the scope of our hypotheses, we may well want to consider causally heterogeneous patchwork hypotheses, as per Bennett's expectation (this symposium) that "in social life there are few simple hypotheses with broad

12 For example, values of some socioeconomic indices may not have been measured or reported with sufficient precision to determine whether a country has crossed specified thresholds.

13 Scope conditions involving categories like "developed countries" or "social democracies" are not probabilistically uncertain but rather semantically fuzzy, until the associated concepts are defined more precisely.

14 Fairfield and Charman 2022, chap. 5 provides guidance on iteratively adjusting scope conditions; see chap. 10 on iterative research.

15 We might imagine inputting the patient-reported symptoms, case history, and results of physical examination and diagnostic tests into some sort of logistic regression model, or neural network, etc., in order to generate a posterior predictive probability distribution over possible diagnoses. Any reasonable model should of course make use of suitable priors or "base-rates" appropriate to the medically relevant reference classes to which the individual belongs, as well as more case-specific information as it becomes available.

16 See also Fairfield and Charman 2022, appendix 12.D.

scope."[17]

We can now address one of Jacobs' central concerns about how the distribution of evidence across cases matters for updating. Referencing the democratic mobilization example, he worries that when aggregating weights of evidence, our approach allows "no distinction to be made between observing ... three highly probative ... pieces of evidence in favor of the communal elites hypothesis (relative to its rivals) within a single case, on the one hand, and observing ... three highly probative pieces spread across three separate cases, on the other hand" (this symposium). If we are comparing two causally uniform hypotheses, $H_{CE}$ = *Slater's communal elites causal logic operates throughout Southeast Asia* vs. $H_{ED}$ = *Economic decline sparks democratic mobilization throughout Southeast Asia*,[18] then indeed it does not matter whether the evidence comes from one case or is spread across three cases, because under either theory, the mechanism is asserted to be the same across all Southeast Asian cases. However, suppose that we compare $H_{CE}$ to a more complex, causally heterogeneous hypothesis $H_{CE/ED}$ = *The communal-elites causal logic operates in the Philippines and Vietnam, whereas economic decline instead sparks democratic mobilization elsewhere in Southeast Asia*. Finding three pieces of evidence from the Philippines that strongly support the communal elites causal logic over the economic decline causal logic would then fail to discriminate between $H_{CE}$ and $H_{CE/ED}$, whereas if one piece of evidence of similar strength for the communal-elites causal logic were found in each of three cases—the Philippines, Vietnam, and Burma, that evidence would support $H_{CE}$ over $H_{CE/ED}$—albeit with only the evidence from Burma contributing inferential weight. If instead the third piece of evidence from Burma favored the economic decline causal logic over the communal elites logic, then the three pieces of evidence taken together support $H_{CE/ED}$ over $H_{CE}$—we might then say that based on our knowledge so far, the communal elites *causal logic* does not generalize beyond the Philippines and Vietnam, but $H_{CE/ED}$ nevertheless provides a viable (if tentative) explanation for democratic mobilization in all Southeast Asian countries.

Accordingly, it is important to remember that whether and how evidence in one case is informative about other cases depends on the hypotheses under consideration. If our inference from the Philippines involves hypotheses scoped to include only this country, then however strongly the posterior odds favor one or the other explanation, those hypotheses would make no predictions whatsoever about what we ought to find in Burma. If we then *tentatively* expand the scope conditions to include all of Southeast Asia, the hypotheses now do make predictions about how things should work in Burma, and *if* these are the only hypotheses under consideration, then evidence collected from the Philippines will indeed shape our current views about the leading hypothesis for understanding democratic mobilization not only in that country, but also in Burma. But if we include hypotheses that postulate operation of different mechanisms in the Philippines and Burma, then it will become important to also look at evidence from Burma.

Generalization then does not happen automatically or by fiat in our approach—we do not get something for nothing, as Jacobs fears. Instead, generalization involves hypothesizing and testing—we *compare* rival hypotheses that make predictions for some shared set of cases. Any background knowledge we have about homogeneity of cases should inform how we craft hypotheses and evaluate their prior odds; updating will depend on what evidence materializes and how and where the predictions of our rival hypotheses diverge.[19]

Some of the skepticism that Jacobs and others have expressed about our approach to generalization may stem from not fully appreciating the conditional and contingent nature of our Bayesian reasoning.[20] We cannot emphasize enough that both our theories and the credences we hold in them are provisional. We are always free to revise hypotheses, whether by changing the causal logic or altering the scope. As we consider a broader set

17 A caveat related to Occam's razor (Fairfield and Charman 2022, chap. 6) applies here. By any sensible measure of complexity, there will be exponentially more complex theories than simple ones that might in principle be considered. This has important consequences. Even if we put more prior probability on the class of complex hypotheses than the class of simple ones, any one complex hypothesis would tend to have lower prior probability than any one simple hypothesis, because there are so many more possibilities of the former class compared to the latter. Accordingly, our best strategy is to start by considering simpler theories, only adding complications or elaborations as necessary, as the simpler theories falter. And by reflecting on how simpler theories fail, we often find hints about how to improve them.

18 See Slater (2009) and Fairfield and Charman 2022, chap. 5.

19 Both Jacobs and Bennett appear to want to presume causal heterogeneity unless there is positive evidence otherwise. But Occam's razor suggests the opposite strategy. As for "building the researcher's beliefs about heterogeneity directly into the likelihoods of the evidence" (Jacobs, this symposium), we contend that this is a job for theory. A carefully articulated and scoped hypothesis builds conjectures about the homogeneity or heterogeneity of cases into the likelihoods of possible evidence.

20 We suspect that some of Jacobs' concerns may also reflect a commitment to working with a potential-outcomes framework and assigning cases to latent causal types (Humphreys & Jacobs 2015), whereas our approach focusses directly on causal explanations articulated in rival hypotheses (see Fairfield and Charman 2022, chap. 9, 395-96). We look forward to further discussing the distinctions between our approaches in a future setting.

of cases, beyond proposing patchwork hypotheses, we might devise new hypotheses that endogenize what we previously considered to be a binary scope condition, so that it becomes part of the (possibly probabilistic) causal logic itself, perhaps as a (no longer binary) moderating variable (Fairfield and Charman 2022, chap. 5, 204-17). After each iteration of hypothesis refinement, we apply the Bayesian apparatus to ask which hypothesis among a set of comparably scoped alternatives provides the best explanation in light of the evidence in hand so far. As data accumulate, a given explanation may gain or lose plausibility in relation to rivals that might posit different or more heterogeneous explanations.

## A Unifying Logic for Inference

We are happy to embrace Soifer's characterization of our book as a Bayesian version of KKV (King, Keohane, and Verba 1994)— we share a similar overarching goal of providing a unified approach to inference that applies to both qualitative and quantitative data (Fairfield and Charman 2022, chap. 9), and "observable implications of theories" do indeed play a central role in our framework. We would characterize KKV as a frequentist-inspired perspective on qualitative research, which we find problematic because according to its own foundational principles, frequentism can only be used to analyze stochastic data. Bayesianism is the only natural and logically rigorous inferential framework that can accommodate both qualitative and quantitative evidence—regardless of what type of information is in hand, inference proceeds according to the same underlying principle: evaluate relative likelihoods for the evidence under rival hypotheses. As such, what Soifer characterizes as a purely practical or "omnivorous" approach for using any informative data to test theories actually rests on deep foundational principles (Fairfield and Charman 2022, chap. 2). Furthermore, Bayesianism is ideally suited for addressing KKV's (King, Keohane, and Verba 1994, 32) central critique of qualitative political science: "the pervasive failure to provide reasonable estimates of the uncertainty of the investigator's inferences." Bayesian probability is an extension of Boolean logic to contexts of uncertainty and limited information; inferences are expressed as posterior odds that characterize how much confidence we hold in a hypothesis relative to rivals given the evidence in hand, or equivalently, how much uncertainty remains regarding which hypothesis provides the best explanation.[21] A Bayesian framework also clarifies that what matters is not how many empirical observations line up with our theory, but rather the *relative likelihood* of the evidence under rival theories.[22]

While we are indeed pushing back on the now conventional QMMR understanding that qualitative and quantitative research follow different logics of inference, we agree with Goertz and Mahoney that these research communities have been characterized by different cultures of inference. But when comparing conventional quantitative research to in-depth qualitative research, we would argue that the cultural difference is marked by frequentism versus intuitive Bayesianism. We suspect that this epistemological mismatch (even if not explicitly recognized as such at the time) is what motivated much of the reaction from qualitative scholars against KKV's prescriptions, some of which impose impractical constraints that are not necessary within a Bayesian framework, including the stricture of testing theory with new data that was not used to inspire or refine the theory.[23] Bayesianism by contrast gives a solid mathematical foundation for iterating between data collection and theory refinement, as well as many other common practices in qualitative research that are not justifiable within a frequentist framework.[24] We hope that qualitative scholars will find these foundations empowering. Bayesian updating in our experience mirrors the way many scholars naturally approach research. And putting our approach into practice involves very little math. Even for those who choose to use the quantified version of Bayes' rule with log-odds, nothing more than addition and subtraction is required.

As to Soifer's perception of a tension between our commitment to qualitative research and our premise that Bayesianism provides a unified inferential framework, we instead see these matters as closely related and complementary. Recognizing Bayesian probability as a universally applicable framework places qualitative evidence on much more equal ground relative to quantitative data and experimental evidence and should thereby help to clarify and revalue the contribution of qualitative information to causal inference, which we understand as inference to best explanation. As discussed in Section 2, Bayesianism imposes no constraints on the notion of causation that hypotheses embrace, so we do not anticipate the particular discord that Soifer contemplates, although we of course recognize that

21 Interestingly, KKV (King, Keohane, and Verba 1994, 32) end their section on "Reporting Uncertainty" by encouraging one to ask: "How much ... would you wager" or "At what odds," which is inherently Bayesian.

22 An additional distinction is that in contrast to frequentist requisites, in a Bayesian framework, all observable implications need not be listed in advance of data collection (Fairfield and Charman 2022, chap. 10).

23 See for example Ragin (1997, 3).

24 See especially Fairfield and Charman 2022, chaps. 10, 12.

adopting a Bayesian framework would require many quantitative scholars as well as many qualitative scholars to change their research practices.

Bouchat expresses more substantial doubts about the value of Bayesianism for equalizing the role of qualitative and quantitative research, based on a claim that quantitative critiques of qualitative research now rest not on issues of inference, but rather on issues of data selection, case selection, generalizability, informativeness of data, and "data validity" or "data quality." First, we fail to see how "data validity" in the sense of measurement validity could underpin quantitative critiques of qualitative research, since if anything qualitative scholars would seem to have the advantage on this front, thanks to in-depth case knowledge.[25] Second, and most importantly, we would counter that none of the other considerations can be separated from inference. Scholars critique these aspects of research because they matter for inference, but how and to what extent they matter depends on the espoused methodology. Frequentism and Bayesianism treat case selection and other aspects of research design very differently. They handle bias differently. They understand and conduct generalization differently. And they evaluate informativeness of data differently. Judgements about "data quality" are likewise directly linked with the ability to draw reliable inferences, so this too is ultimately a methodological question. To focus on any of the particular concerns Bouchat mentions while overlooking methodological distinctions between frequentism (the dominant framework for quantitative political science) and Bayesianism, or any other inferential approach that qualitative scholars have espoused, is to miss the underlying source of disagreements and tension. Clarifying these distinctions is a central aim of our book.

Bouchat (this symposium) goes on to say that the desired goal should be to "treat evidence derived qualitatively as equal with that measured and collected quantitatively"—but we contend that Bayesianism does just that, precisely by virtue of drawing inferences from all data in same manner. Of course, not all evidence will be equal in terms of its inferential import, but inferential import depends on how strongly the evidence in hand discriminates between rival hypotheses, not whether it is qualitative or quantitative. As for the assertion that Bayesianism "does not at all resolve ... what qualifies as good data," we are perplexed, considering that Bayesianism to our minds provides a clear and straightforward answer: "good" data are *informative* data, namely, any observations that are more expected under one hypothesis compared to rivals. The more divergent the likelihood of the data under rival hypotheses, the more informative the data, and hence the "better" the data. If the data are noisy or imperfect, then Bayesians can and should take that into account. If there is a question about the validity of a measurement (i.e., whether it captures the concept or variable of interest), Bayesians can and should take that into account as well, by conditioning on the raw data as they are, not on the value of some variable that we hoped to measure but did not. So again, we do not see concerns about data validity or data quality as either a fundamental source of difference between quantitative and qualitative research, or as a challenge to our argument that Bayesianism serves as a universal framework for inference that revalues the contribution of qualitative evidence.

A second leg of Bouchat's critique, in our perhaps imperfect understanding, is that by espousing an objective Bayesian framework rather than fully embracing subjectivity, we necessarily undermine qualitative research, which inherently involves subjectivity. We find this reading counter to the intent and substance of our book; we of course fully agree that qualitative research "does not lack in its scientific value by leveraging data or insights that defy easy quantification" (this symposium). Throughout, we acknowledge that subjective inferences are necessary in practice, and we emphasize that quantitative social science is no exception—not only to the extent that it draws on qualitative information that has been imperfectly quantified to construct datasets, but also through the many decisions made when elaborating models that necessarily require scholarly judgement.[26] But our goal is articulate principles and illustrate practices that can help social scientists to reason as rationally and objectively as possible about the way the world works. Understanding and following Bayesian principles helps our subjective judgements better approximate the ideal of rational inference, while simultaneously allowing us to leverage all the information in nuanced, detailed, qualitative evidence. We do not think that working to minimize subjectivity in inference cedes ground to any claims that quantitative research is superior due to greater objectivity—again, we explicitly argue against the notion that objectivity vs. subjectivity distinguishes quantitative from qualitative social science (Fairfield and Charman 2022, chap. 9, 440-45)—or that it undermines our vision for Bayesanism as a unified inferential framework. We

25 Bouchat (this symposium) further mentions validity "in the substantive sense" of "studies that do not identify causation"—we are not sure exactly what is meant here, whether it be all qualitative research, which by frequentist quantitative standards cannot produce causal identification, or specifically qualitative research that does not focus on explanation. But it is worth emphasizing that our book primarily speaks to qualitative research that aims to make causal claims.

26 We return to this point in Section 5.

return in Section 6 to clarify some specific points about priors and knowledge accumulation that might have fostered perceptions to the contrary.

## Formalization and Analytical Explicitness

While our approach to inference is guided by the formal apparatus of probability theory, we do not, as Jacobs notes, formalize the derivation of likelihood ratios. Formalization would require devising a statistical model (e.g., regression-like structural equations, input-output tables, or an instantiation of a DAG) capable of producing precise numerical likelihoods in an algorithmic way for every possible piece of evidence that might be observed during data collection. Instead, we quantify relative likelihoods only for the empirical observations that do turn up, once they are in hand, based on informal but careful verbal reasoning about the predictions that our plain-language hypotheses suggest.[27]

We take Jacobs' point that formalization and objectivity are conceptually distinct, and we recognize that describing formalization as creating a "veneer of objectivity" may not have adequately conveyed why we prefer to reason informally about likelihood ratios. To clarify, we contend that formalization of the sort Jacobs has in mind is essentially impossible when working with the kind of detailed qualitative evidence that is the central concern of our book—that is, we do see "fundamental limits" to formalization in this context.

The problem lies in that formalization requires specifying probabilities for all possible empirical observations in advance, but we cannot hope to even envision all such possibilities when the evidence in question involves open-ended responses from expert informants, passages from archival sources, accounts from newspapers, firsthand observations of human behavior, visual information from campaign ads, and so forth. To illustrate the scale of the problem, consider information that a scholar might elicit from an expert informant during an interview. If the informant gives a three-sentence reply to just a single question, there may be on the order of $10^{200}$ possible responses (taking into account the average length of a sentence in English and ignoring non-verbal cues) that would need to be enumerated and then assigned likelihoods.

Any effort to formalize hypothesis testing with this kind of qualitative evidence would require massive coarse-graining of potential observations into a manageable number of categories. But details in the evidence (e.g., tone of voice, body language, identity of the informant, context in which remarks are made) can matter greatly for likelihoods, so the probability that the model assigns to any one of the coarse-grained evidentiary types it specifies may not be an adequate approximation for the likelihood of any concrete qualitative empirical observation that turns up. The coarse-graining required for formalization in essence throws away relevant information in the evidence and distorts the conclusions. In contrast, our approach avoids what we see as the unnecessary and near impossible effort of assigning probabilities to the myriad possible evidentiary observations that might have materialized but did not, while allowing us to use all the information in our evidence.[28]

Subjectivity inevitably enters our informal approach when reasoning about which of one or more rival hypotheses (expressed in ordinary language) makes an evidentiary observation more expected, and in assigning numerical values to represent our judgements about evidentiary weight.[29] In logical Bayesianism, objective probabilities are determined exclusively by the information available. Subjective probabilities draw not only on the information available, but also on judgement, which should be informed by expertise and experience, but will also involve some degree of arbitrariness. While the guidance in our book aims to help subjective probabilities better approximate the logical Bayeisan ideal, the fact remains that there is no strictly objective way to quantify probabilities for complex, nuanced, inherently qualitative information about the socio-political world.

But we maintain that our approach makes this subjectivity transparent and invites discussion among scholars who may think differently, which in turn facilitates consensus building, or at least clarification of where any why scholars disagree. As such, we would say that we achieve the same goals without a formal model that Jacobs highlights in writing that "a model representing the researcher's beliefs about how the world works, and from which the likelihoods are then derived, makes explicit elements of the analysis that will otherwise remain implicit." In our approach, assigning some qualitative observation a weight of 10 dB in favor of $H_1$ vs. $H_2$ clearly conveys our degrees of belief to readers. And we accompany this quantitative judgement up front with a written explanation for why we consider the evidence to be moderately more expected under $H_1$ vs. $H_2$. In contrast, the formalized approach advocated by Jacobs to our minds hides the researcher's views

27 We are referring here to our approach for log-odds updating (Fairfield and Charman 2022, chap. 4), which we call "explicit Bayesian analysis" (for lack of a better term), as opposed to "heuristic Bayesian reasoning" (Fairfield and Charman 2022, chap. 3), which applies the same thought process but stops short of quantification.

28 See Fairfield and Charman 2022, chap. 10, 479, 495-98.

29 Subjectivity of course also enters when evaluating priors (see Section 6).

within the intricacies of the model, in a way that makes it more difficult for readers to understand, evaluate, and critique—at least when the evidence involves inherently qualitative information. When applying our approach, if another scholar asks why we deemed the weight of evidence to be 10 dB rather than 15 dB, we can have a conversation on the spot, which may lead us to better articulate our reasoning, specify our hypotheses more clearly, or revise our views. When employing formal models of the sort proposed by Humphreys and Jacobs (2023), an analogous discussion would involve questions about distributions over latent variables or parameters, the precise form of structural equations, or other highly technical attributes of the model that are more difficult to connect to the substantive meaning of a theory and the evidence in hand.

We of course do not object to formalization in all contexts. But for in-depth qualitative research, formalization would involve replacing a manageable number of subjective but direct judgements about likelihood ratios for observed evidence with a vast number of ultimately subjective choices about technical intricacies of the model.[30] We envision few benefits in terms of explicitness or transparency to embedding probabilistic judgements in multiple layers of parameterizations with limited interpretability when the inferences we care about are the relative plausibilities of rival theories that provide distinct explanations for socio-political phenomena. This is what we had in mind when writing that formalization "simply pushes the subjectivity back deeper into the model" (Fairfield and Charman 2022, 442).

Lastly, we do not agree with Jacobs' remark that "formally deriving priors and likelihoods from a single underlying model forces internal consistency among the inputs to Bayesian analysis," (this symposium) for the simple reason that a model itself cannot provide all of its own priors. For DAGs of the sort discussed in Humphreys and Jacobs (2023), each node in the graphical model will typically require many *exogenous* inputs determining prior probabilities over various nodal types. More generally, hierarchical modeling can push the exogenous probability inputs into deeper layers, but that does not circumvent the need to make largely arbitrary choices about prior distributions for hyper-parameters which may influence the observable predictions of the model in ways that are difficult to discern. And what we consider the most important prior probabilities for theory testing, namely those specifying relative plausibilities for the overarching model families that offer competing explanatory frameworks (e.g., specifying distinct DAG topologies), can never be regarded as part of the model—hypotheses cannot assert their own degree of plausibility.

We believe our differences of perspective on these points stem from the distinct research contexts we focus on as well as our orientation toward hypothesis testing. As already discussed, our work focusses on analyzing open-ended, detailed qualitative observations, whereas Humphreys and Jacobs (2023) apply their formalized approach primarily to moderate numbers of variables that assume only a moderate number of values. Furthermore, we engage in theory testing by comparing rival hypotheses, which would be the heuristic or qualitative analog of comparing distinct model families, whereas Humphreys and Jacobs largely focus on what would be considered parameter estimation in standard statistical parlance, along with other inferences within a single chosen model family.[31]

## Priors and Knowledge Accumulation

While the role of prior probabilities in Bayesian inference does constitute a major departure from frequentist frameworks, the importance of priors is sometimes exaggerated by critics. In our view, Bouchat's claim (this symposium) that the guidance our book offers for Bayesian reasoning "only makes sense conditional on the establishment of prior probabilities" is similarly exaggerated. We instead hold that weights of evidence merit much greater attention than priors in qualitative research that aims to bring new evidence to light.[32] Moreover, the same guidance provided in our book would directly apply to research agendas that aim to systematically construct informed priors or characterize the existing state of knowledge in a field. Before elaborating these points, we briefly review our approach to priors.

Recall that probabilities within objective, or logical, Bayesianism are degrees belief determined by states of

30 In Humphreys and Jacobs' (2023) potential-outcomes framework, the total number of causal types, and the parameterizations associated therewith, grow super-exponentially with the number of distinct values or categories that the independent and dependent variables can assume. While the growth in complexity can be partly tamed by a choice of a particular DAG topology, full formalization will unavoidably require an enormous number of largely subjective decisions to give concrete shape to the probability model.

31 Translated into Humpheys and Jacobs' (2023) framework, what we are doing would involve comparing distinct DAG topologies involving substantially different nodes or different connections between nodes.

32 Priors matter more for quantitative research involving parameterized models. Here we are interested in prior odds on competing theories (or model families).

knowledge.[33] Accordingly, the aspirational goal would be to incorporate all relevant initial information (and nothing else) into our prior odds. In principle, we would go back to a state of minimal knowledge or ignorance that justifies equal odds on hypotheses of comparable complexity, and then build up to our present state of knowledge by employing Bayes' rule, effectively incorporating all of our initial information as evidence (Fairfield and Charman 2022, chap. 3, 96). Practically speaking, however, we usually have too much initial knowledge to carry out this procedure, short of turning the construction of priors into the sole focus of research. For work that aims to bring new evidence to light, we will have to make do with subjective approximations to the logical Bayesian ideal, in that we will need to use judgement to guide us rather than attempting a full and systematic accounting of background information. As such, we suggest two options: (1) articulate informed priors as best as possible, explaining how key elements of background knowledge motivate these judgments, or (2) just start from equal odds on the salient hypotheses, which focusses attention on the evidence at hand and in essence allows readers to supply their own priors. Whether starting with informed priors or indifference priors, it is sensible and straightforward to conduct sensitivity analysis by exploring the import of different priors, including priors that anticipate the reaction of skeptical readers whose background knowledge might lead them to prefer a rival hypothesis over the author's favored argument. Such sensitivity analysis is almost trivial when working with the log-odds form of Bayes' rule.

We now turn to clarifying several points with regard to Bouchat's critique. First, indifference priors in qualitative Bayesian reasoning are not meant to "reflect impartiality and objectivity." As explicated above, true objectivity would involve systematically incorporating every relevant element of the scholar's background knowledge into their prior odds, which as a practical matter may be impossible in most social science contexts because we simply possess too much background knowledge. Instead, using indifference priors in contexts where we do not actually find ourselves in an initial state of ignorance is a pragmatic recommendation to address the reality that readers will inevitably bring their own very different priors, based on very different background information, to bare on our work. Given this reality—and stressing how dramatically background knowledge and hence prior beliefs can vary among scholars—we contend that the most important task is to focus on the inferential weight of the evidence we are contributing to the literature. The greater the weight of evidence in hand, the less priors will matter for posterior judgements, and scholars who start with different priors may still end up favoring the same hypothesis in light of the evidence. And even if priors remain poorly specified or contested, carefully analyzing the weight of the evidence in hand can still make an important contribution to knowledge accumulation. Furthermore, by reporting weights of evidence, or equivalently, posterior log-odds based on indifference priors, authors and readers can immediately discern what strength of prior belief would be needed to overcome the import of the new evidence.[34]

Second, we have no objections to subjective priors as a heuristic, so long as they aim to reflect the scholar's empirical background knowledge, rather than desires about how the world ought to work or empirically unjustified preferences for a pet theory—these kinds of considerations are subjective in a non-scientific sense, as opposed to subjective in the sense of varying across individuals who simply possess different information.[35] Values, desires, and personal preferences can certainly guide the choice of research questions and ethical research practices, but they should not affect inferences from empirical evidence.[36]

As for the alternative of eliciting priors and pooling opinions from experts, we recognize that this is an active area of scholarship within subjective Bayesianism, and

33 Strictly speaking, probabilities cannot be "measured," since they are epistemological rather than empirical. Bouchat's comments, however, frequently refer to measurement, in ways that leave us unsure of the intended meaning. For example, Bouchat reads us as advocating that scholars "conform as much as possible to objective aims, measurement, and likelihood specification." This characterization is correct on the first and last accounts, but measurement in this context does not comport. Our book is not about measurement; it is primarily about inference with qualitative evidence, which involves assessing likelihood ratios rather than measuring or scoring some variable or concept.

34 Notice that using indifference priors is mathematically identical to dropping priors and simply reporting weights of evidence, which is a common Bayesian convention.

35 The pages Bouchat highlights as evidencing our outsized "concern about undue influence, subjectivity, and bias" in priors involve our response to particular issues that *other scholars* have posed to us, specifically regarding our treatment of iterative research. We view their concerns as legitimate but easily addressed (Fairfield and Charman 2022, chap. 10, 491-92).

36 Regarding Bouchat's interest in how to rigorously establish informed priors, beyond our guidance to provide an explanation for one's view that highlights the most salient elements of one's background knowledge, our advice would be to identify and analyze concrete empirical information from the literature reviewed, and then build up from ignorance priors to informed priors using Bayes' rule. Short of conducting a Bayesian meta-analysis (see below), which would be a research project unto itself, there is no "off the shelf" instruction manual for how to circumscribe this process to make it a feasible task—scholars would have to exercise judgement and explain their decisions, just as they do when arriving at informed priors as per our more informal guidelines.

we acknowledge that this kind of approach may be useful for some research agendas. However, there is no widely accepted algorithm for these tasks that can be fully justified with objective Bayesian principles. We also caution that even a rigorous methodology for aggregating potentially divergent expert opinions may run up against the limitation of experts who are not themselves Bayesian reasoners. While expert opinions draw on expert knowledge, experts may not arrive at their opinions via any sort of coherent Bayesian principles. And it is far from clear whether imperfections in individual scholars' reasoning can be averaged away through the aggregation process, especially if "conventional wisdom" leads to positively correlated errors.

From a logical or objective Bayesian perspective, we would ideally want to pool experts' *empirical knowledge*, rather than experts' opinions, and then carefully analyze that knowledge to arrive at relative odds on salient hypotheses. At least in principle, this could be done by training experts in Bayesian reasoning and holding workshops where knowledge is shared, analyzed, and debated (along the lines of the research agenda Bennett mentions). Importantly, notice that this process would involve treating what would otherwise be background information as evidence, and would thus become identical to assessing and scrutinizing weights of evidence as per the guidelines in our book, with a focus on known facts within a research community rather than new evidence obtained through original research.[37]

Turning to knowledge accumulation, Bayesianism is an ideal framework for learning both across different components of a single study and across distinct studies. Whatever the data source or type, weights of evidence accumulate additively,[38] and prior log-odds add to the total weight of evidence to yield posterior log-odds. In the first context, scholars conducting, for example, Bayesian analysis of a quantitative dataset followed by case studies (or vice versa), can employ their posteriors from the first component of research as their priors for the second component of research. In this manner, knowledge accumulates naturally across aspects of the research that draw on distinct kinds of data, without recourse to different methodologies that draw on incompatible epistemologies and produce findings that are not easily integrated. Here we are not sure what to make of Bouchat's suggestion that our recommendation for scholars to use their own background knowledge and priors undermines our vision for Bayesianism as a unified inferential framework, considering that learning across components of a study proceeds in the manner described above regardless of how priors for the first component of research were generated. If the priors and weights of evidence are reported separately, then readers can substitute their own priors, or if they wish, try to formulate some sort of consensus prior for the relevant research community.

Regarding knowledge accumulation more broadly—not just across different components of a research project, but across distinct studies, perhaps aiming to draw on all relevant published literature—one enters the realm of what we might call meta-analysis. While this is not the focus of our book, the same principles and guidance apply at this level. Bayesian macroknowledge building or meta-analysis would be straightforward if all studies in the literature reported Bayesian weights of evidence with respect to leading rival explanations: weights of evidence would then be additive across studies in the same way that they are additive within studies.[39] But reporting weights of evidence has not been standard practice in social science.[40]

Given this status quo, a careful meta-analysis designed to assess the state of knowledge in a field would require (1) devising a common set of explanatory hypotheses to compare that includes the leading arguments under debate, (2) extracting concrete empirical evidence from literature in the field, and (3) conducting Bayesian inference. While we see ample potential here for major contributions to social science, this kind of project would involve a very substantial amount of effort. For qualitative research on, say, state building, one would need to employ a team of trained scholars, and ideally engage experts in a process of scrutiny, adjudication, and consensus building.[41] If a project of this sort proved achievable, scholars could then employ the resulting posteriors as priors for additional research on the topic. But significant challenges remain, in that once someone invents a new hypothesis to test, rigorously speaking, they would have to go back through the entire body of evidence considered in the meta-analysis to construct prior log-odds for the new hypotheses relative to rivals.[42]

37 Alternatively, one could interview or poll domain experts about competing theories and try to use these responses as *testimonial* evidence, but the likelihoods would be extremely challenging to assess.

38 Provided that allowance is made for possible logical dependency in the data given the hypotheses.

39 Again, modulo any dependency considerations. And additional analysis, revisiting the original data, would of course have to be conducted for new hypotheses that were not previously assessed.

40 An obvious first step is to train scholars in Bayesian reasoning, which is the purpose of our book.

41 We have been actively looking into opportunities for conducting precisely this kind of research and would be happy to hear from any interested potential collaborators.

42 Fairfield and Charman 2022, chap. 10.

If a scholar's primary goal is to contribute new evidence to the debate, then we would reiterate the advice in our book: rather than undertaking the mountain of effort needed to systematically incorporate all relevant background knowledge from existing literature, articulate priors that aim to reflect the most consequential elements of your own background knowledge, and then focus on evaluating new evidence.

As for accumulating knowledge across disciplines, the same principles expounded in our book are directly applicable here too—Bayesianism is a natural framework for knowledge accumulation in all contexts. Our current research on covid origins demonstrates how an informal Bayesian framework can be used to organize and analyze diverse kinds of evidence produced by multiple fields of inquiry, ranging from genomic information and epidemiological evidence to information from observational field work, testimonial accounts, and journalistic reports. Our research has involved reviewing literature and interviewing expert informants across disciplines as diverse as virology, genomics, zoology, medicine, geography, and political science. The same caveats expounded in our book apply in this context as well. Quantifying weights of evidence is an undeniable challenge—whether the evidence involves readily quantifiable data about the spatial location of early covid cases, or qualitative observations that coronavirus research at the Wuhan Institute of Virology was conducted at relatively low laboratory bio-safety levels. And there may well be more arbitrariness in some of our weights of evidence than in others; as per the passage Bouchart highlights (Fairfield and Charman 2022, 444-45), we need to view our quantification efforts with some healthy skepticism, keeping in mind that our judgements are provisional and subject to revision. The imprecision of our weights of evidence can be partly addressed through sensitivity analysis (277, 280-82)—we specify a range of values for each piece of evidence rather than reporting only a single number. But more importantly, our estimates could serve as a starting point for structured scrutiny and debate among experts.[43]

Notwithstanding the limitation that many kinds of information do not yield objectively quantifiable probabilities, we view (approximately) objective Bayesianism as the only natural framework for knowledge accumulation, especially when it comes to learning across diverse kinds or sources of evidence. Frequentism in principle rejects the use of any data that are not generated by some stochastic process, and because probabilities cannot be assigned to theories, frequentist-based approaches are awkward at best when it comes to combining evidence or conclusions across multiple studies.[44] Fully subjective Bayesianism allows supposedly rational agents who have exactly the same information to come to different probabilistic conclusions, with no way to reconcile the discrepancy, so it is not even clear what knowledge accumulation should mean in this context. For complex and controversial cases like covid origins, the Bayesian approach offers additional benefits—it forces us to take seriously rival explanations that may run counter to what we want to be true or what we initially think is true, and it can reveal where and why reasoning about key pieces of evidence among the public, in the press, and even in peer-reviewed literature may go wrong.

## Conclusion

Soifer's remarks contemplate how scholars of different persuasions will react to our work—in our experience to date, enthusiasts and skeptics have not been split along traditional quantitative vs. qualitative methodological divides. We take that as a positive sign, considering that our goal was not to write a book that everyone would agree with and readily adopt, but rather to shake up existing divisions within the discipline, rechart the methodological landscape, and challenge scholars to rethink which of their practices are justified and valuable, and which could be improved to yield more reliable and consistent inferences. We would say the more Bayes the better to that end, but to Soifer's query, we grant that readers who are reluctant to embrace the full Bayesian apparatus can still benefit from incorporating some of the lessons of Bayesian reasoning into their work. We thank the discussants again and welcome further debate moving forward.

43 Unfortunately, we have found that this particular question has become so polarized and politicized that few experts have been willing to engage in this fashion. We would also like to note here that while some political scientists have expressed trepidation about quantifying degrees of belief when working with qualitative evidence, in our view, the benefits for promoting consistency of reasoning across multiple pieces of evidence and systematically aggregating their inferential import can outweigh concerns about false precision—particularly in contexts where the evidentiary observations do not all tilt the balance in favor of the same hypothesis (Fairfield and Charman 2022, chap. 4). In these situations, drawing conclusions requires going beyond qualitative judgements about individual weights of evidence. We will have to ask, for example, whether two pieces of evidence that each moderately favor $H_1$ over $H_2$ together outweigh one piece of evidence that strongly favors $H_2$ over $H_1$. In making a judgement, we are at least implicitly moving toward quantification, and explicitly quantifying makes our decisions more transparent. If desired, one could always translate the aggregate quantified weight of evidence back into a qualitative description (e.g., weak, moderate, strong, very strong...) to avoid conveying false precision.

44 See Fairfield and Charman 2022, chap. 8.

## References

Chan, Alina, and Matt Ridley. 2021. *Viral: The Search for the Origin of Covid-19*. Harper Collins.

Fairfield, Tasha, and Andrew E. Charman. 2017. "Explicit Bayesian Analysis for Process Tracing." *Political Analysis* 25 no. 3 (July): 363–80.

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

Humphreys, Macartan, and Alan M. Jacobs. 2023. *Integrated Inferences: Causal Models for Qualitative and Mixed-Method Inference*. New York: Cambridge University Press.

King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.

Ragin, Charles. 1997. "Turning the Tables: How Case-Oriented Research Challenges Variable-Oriented Research." *Comparative Social Research* 16:27–42.

Slater, Dan. 2009. "Revolutions, Crackdowns, and Quiescence: Communal Elites and Democratic Mobilization in Southeast Asia." *American Journal of Sociology* 115(1):203–254.

*Qualitative & Multi-Method Research* 18(2), "Symposium: Authors meet Critics—Veil of ignorance Process Tracing."

*Qualitative & Multi-Method Research* 14(1)), "Symposium: Causal Mechanisms and Process Tracing."

QMMR Qualitative & Multi-Method Research

Notes from the Field

Qualitative and Multi-Method Research Fall 2022, Vol. 20.2 https://doi.org/10.5281/zenodo.8326524

# Boundaries Unsettled: Invisible Threats and Activist Scholarship in Uruguay

Francesca Lessa
*University College London*

February 10, 2017, was a warm summer evening in Mendoza, Argentina. The thick blackout curtains were trying, unsuccessfully, to keep the torrid heat out of the room. In the sunset light, I glanced at my phone on the bedside cabinet. A message from my friend Silvia flashed on the screen.

Although my memories of that hot summer evening are fuzzy in places, I will never forget the content of that WhatsApp message: Silvia wanted me to know that she had heard my name on the evening news in Montevideo, Uruguay, as integrating a death list composed of 13 people, mostly authorities (including the country's attorney general and minister of defense), lawyers and human rights defenders, 10 of whom were Uruguayans and three foreigners. I knew many of them personally given the research I had been conducting on impunity for dictatorship-era crimes in Uruguay for almost ten years.

For the next few hours, I was in a shock-like state trying to make sense of what was unfolding.

Me? On a death list? In Uruguay?

I did not tell anyone about the death threats for the first 24 hours: I was unable to find the words to articulate the situation, which seemed rather surreal in those initial moments. Nothing in all the training courses I had completed as a researcher in my years at the University of Oxford—on fieldwork security, risk assessment, ethics, and vicarious trauma—could have prepared me for this.

A previously unknown group in Uruguay had disseminated the death list to the media, local authorities, and also emailed it directly some of the threatened people themselves. I had not received anything, though, aside from Silvia's message. The death threats came from the self-proclaimed "Comando General Pedro Barneix" and read as follows (IACHR 2017):

"The suicide of General Pedro Barneix will not remain unpunished… No more suicides or unjust prosecutions will be accepted. From now on, for every suicide we will kill three people selected at random from the following list."

The communique then listed thirteen names and ended with an ominous warning: "And we have several more, whose addresses and habits we have already compiled."[1]

The group named itself in homage to Pedro Barneix, a retired Uruguayan general who had been indicted for the murder of ice-cream maker and left-wing sympathizer Aldo Perrini, in the city of Colonia in 1974, during the country's 1973 to 1985 military dictatorship. On September 2, 2015, when the police went to his house to formally notify Barneix of his pre-trial detention, he killed himself. The general had been a trustworthy associate of President Tabaré Vázquez, of the left-wing coalition "Frente Amplio" (Broad Front), during his first mandate between 2005 and 2010. Vázquez had in fact appointed Barneix and General Carlos Díaz in 2005 to participate in an investigative commission within the Army to gather information on the fate of the disappeared (El País 2017).

In the next few pages, I reflect upon the experience of receiving death threats whilst on extended fieldwork in South America, the challenges I faced both personally and professionally as a result, how I dealt with them, and how that experience has shaped the relationship with my research communities. This article is written in an autoethnographic style that includes emotions and turning points, as well as "interpretation, reflection, and direct experience, which shows vulnerability rather than distance" (Carspecken 2023, 3).

## Fieldwork Under Threat

In the following days after receiving Silvia's message, I tried to determine from Mendoza the contours of what was exactly unfolding in Uruguay. It was not until two weeks later that, on February 24, I eventually received—in response to my inquiries—an email from the General Directorate of Information and Intelligence of the Uruguayan police that officially informed me that my name was "effectively" included amongst those that the Comando Barneix had threatened with death. With this official confirmation in my hand, I proceeded to inform my line managers and braced myself for the oncoming storm.

The existing literature on research methods and ethics does not contemplate nor discuss the challenges that I faced, both personally and professionally, because of these death threats. A brief review of the scholarship finds numerous publications on conducting fieldwork in risky and violent contexts (Nordstrom and Robben 1996; Sriram et al. 2009; Mac Ginty, Brett, and Vogel 2021; Schultz 2021), as well as on researchers' positionality and reflexivity (Kohl and McCutcheon 2015; Berger 2015; Folkes 2023). There is no discussion, however, of what happens when the researcher becomes part of the dynamics that she or he is studying, when the boundaries become so blurred, overturned, unsettled, when the researcher has turned into the "researched." The closest I could find is the interesting article by Melissa Mendez (2023, 93) who introduces the concept of "victim-as-researcher," to identify people who have been victims of "a violent, physical crime" and have afterwards conducted projects that required them to interview "offenders who have been perpetrators of criminal acts" similar to the crimes they experienced.

This is nonetheless still different to what happened to me. By all accounts, Uruguay is one of the safest countries in South America. Because of this, I had cleared and achieved approval for my risk assessment rather easily: I had conducted research on impunity for dictatorship-era human rights violations in the country since 2007 and undertaken countless trouble-free trips there. By 2016, I also had a large existing network of people and contacts on the ground, which constituted a plus in terms of my risk assessment. What neither myself nor my colleagues in Oxford at the time could have envisioned was that I would be specifically targeted because of the very research that I had been carrying out for almost a decade.

The objectives of the threats were both broad and specific. Broadly, to try to stop—or at least delay—the incipient wave of prosecutions that had finally begun in Uruguay after decades of impunity. The Comando Barneix spoke of "unjust prosecutions" in its email espousing the death threats, and named itself after an ex-general who was, at the time of his suicide, facing trial for murder. Specifically, to silence the voices of numerous people involved in their different capacities in human rights issues relating to the recent past in Uruguay. This included me—an academic who had decided to focus her work on what Uruguayans lovingly call "*el paisito*" (the small country).

Serving as a backdrop to this situation was the tragic fate of Giulio Regeni, an Italian PhD student at the University of Cambridge who had been abducted, tortured, and murdered in January 2016 by intelligence officers of the dictatorship of Abdel Fattah Al-Sisi while conducting field research in Cairo. Because of what had happened to Giulio just a year earlier, neither the University of Oxford nor the Italian Embassy in Uruguay were willing to take any chances. They wanted me to return to Oxford and to Italy, respectively.

Uruguayan authorities did not seem interested in seriously investigating the threats, nor did they offer much protection or support to any of the threatened

1 Author's translation from the original Spanish.

individuals. In this complex scenario, and with the prospect of me having to spend another 18 months in Uruguay, the University of Oxford, the insurance provider, and the Italian Embassy all concurred that I should not return to Uruguay, not even to pack my belongings.

Eventually, I was able to reach a compromise with the University of Oxford: I would relocate to Buenos Aires, where I had previously lived between 2014 and 2016, and continue the project from there. But should anything else happen, I agreed to return to Oxford immediately.

## A Winding Road

In the following weeks, as I tried to salvage my research project under threat, as well as myself, I faced two sets of challenges: one personal, one professional.

Personally, I had become very fond of Uruguay over the years since my first trip there in September 2007. Ten years later, I regularly visited the country not only for research purposes but also because I had developed many connections and friendships. The most difficult aspect for me was accepting that Uruguay, a place where I had felt safe, which I had considered a second home, and what my friend Fernando jokingly said was "*mi lugar en el mundo*," (my place in the world), was so no longer. This loss of certainties was profoundly unsettling.

Professionally, the most urgent challenge was redesigning my project. In some cases, fieldwork does throw the basic premises of a project, such as the research question or case selection mechanism, upside down, a scenario that La Porte (2014, 414) labelled a "crisis of research design." I faced a crisis of research design, of sorts. I had to redesign my project whilst already in the field, but because I had been cut off—for my own safety—from my primary research site and the sources of data (archives, prospective interviewees) that I had intended to use.

After several years of unsuccessful fundraising efforts, in early 2016, I had finally secured a Marie Skłodowska-Curie Global Fellowship, to study the crimes of Operation Condor and probe the response of national justice systems to these transnational atrocities through the lens of Uruguay. Since Uruguayan citizens had been abducted in each of the Condor member states, by reconstructing their cases, I planned to study the whole network of transnational repression and its *modus operandi*. However, I could no longer set foot in Uruguay, at least for the foreseeable future.

Back in Buenos Aires, I grabbed the broken pieces of my original project and faced the task of reorganizing my research plans. At this time, I received the solidarity of numerous peoples and NGOs, which was invaluable to keep me going.

My initial methodology revolved around the combination of three sets of primary sources in Uruguay: archives, legal documents and the monitoring of criminal trials, and interviews. I had to adjust the project so that I could rely on those same sources but from any of the other Condor countries—some of which were unexpected.

When it came to archives, Carlos Osorio of the National Security Archive in Washington and Jair Krischke of the Justice and Human Rights Movement in Porto Alegre, Brazil, both opened the doors of their non-governmental organizations and said I was welcome to use their records instead. Regarding legal documents and trial monitoring, I could no longer follow the Condor trials taking place in Uruguay, but with support from Jorge Ithurburu, president of the Italian NGO 24 marzo, I was able to focus my attention on the trial for Operation Condor crimes in Italy. Underway since 2015, this criminal process probed the murders of 23 Italian citizens, 18 Uruguayans, and two Argentines. In the midst of so much uncertainty, I travelled to Rome a few months later in December 2017, and that trip was like a second chance: I could somehow recover this project and felt I was beginning to do so. As for interviews, since Uruguayans often travel to Buenos Aires for weekends and holidays, I could still interview some of the research participants, who generously donated their time during such trips. Moreover, having expanded my focus to include victims of Operation Condor of all nationalities, I conducted additional interviews in Brazil, Chile, Argentina, the US, and Italy.

I wish I could say that there was a clear strategy and plan of action, as I put back together what felt like a broken project, but it would be a lie: I tried to develop a coherent whole using the pieces I had already gathered and with the new ones I was able to access under my troubling new circumstances.

One unexpected and positive development was the creation of the database on "South America's Transnational Human Rights Violations (1969-1981)." This unique and comprehensive dataset began as a simple excel sheet in which I had listed several names of Uruguayan and Argentine victims of Operation Condor, to provide guidance to my research assistant, Nuria, who was tasked with completing the review of the archives of the Ministry of Foreign Affairs in Montevideo which I had started five months earlier. Ultimately, it became the database it is today due to a collaboration with Argentine sociologist and database expert, Lorena Balardini. The insights that emerged from the analysis of this dataset were instrumental for two reasons. First, they enabled me to develop an original five-phase periodization of transnational repression in South America between

1969 and 1981 that better shows the evolution of the dynamics that led to the emergence and downfall of Operation Condor—which I discuss at length in *The Condor Trials* monograph (Lessa 2022). Second, the data compiled on the 805 victims provided evidence to substantiate new findings, while giving additional weight to existing conclusions on transnational repression in South America. For example, on the one hand, the dataset confirmed that Argentina was the main operative theatre of transnational repression, with 68% of victims being murdered or initially abducted there—a conclusion that had been pointed to by the criminal trials. On the other, it challenged the evidence by US and South American archival documents that justified the emergence of Operation Condor in 1975 as a way to counter the coordination among guerrilla groups, known as the Revolutionary Coordinating Junta (JCR from its Spanish acronym), that had been underway since 1974. The dataset not only showed transnational repression episodes that dated back to 1969—so much earlier than 1974—but also that the majority of victims pursued were in fact political and social activists, not members of the JCR.

## Activist Scholarship

Looking back at the first 12 months following the death threats, I operated as a firefighter that was always on call: I was constantly resolving various crises, whether it was finding a new host institution and supervisor, sorting out the paperwork needed for my visa for Argentina, finding a new place to live, in constant communication with embassies and consulates, dealing with the travel insurance company and their security consultants, and so forth.

Because of the solidarity and support that I received from family, friends, and colleagues on both sides of the Atlantic, I was able to regroup and get all the data I needed in the remaining 18 months in Argentina, and through additional trips to the US, Chile, Brazil, and Italy.

At that time, while I was permanently putting out fires to keep the project going, I did not fully realize a challenge that would become long-lasting: I was no longer a distant observer to the dynamics of impunity that I had scrutinized for a decade in Uruguay, I had become absorbed by my research topic.

To be fair, I had never been "a distant observer" in the sense that, in my opinion, when it comes to issues of human rights violations and injustice, impartiality and objectivity are not feasible. My engagement with local communities potentially did not amount to what anthropology scholars qualify as a "militant ethnographer," but I was at least "a committed scholar," one that produces sympathetic knowledge that is useful to social movements and struggles (Valenzuela-Fuentes 2019, 722). Professor Ken Booth (1997, 115) wonderfully depicts the "special and privileged role" that academics have, through knowledge, "to unsilence the silenced; […] to speak up for those who do not have a voice."

By revealing the policies and politics of impunity in democratic Uruguay, I had exposed the country's failure to comply with the international human rights obligations that it had voluntarily assumed, and to deliver justice to the victims and their families, as well as the broader society, whose rights had been systematically violated under twelve years of state terror. I also brought attention to the fact that impunity was a clear obstacle to putting in place guarantees of non-repetition and, thus, continued to generate conditions whereby human rights would likely be violated again.

My activist scholarship was the result of the profound connections to Uruguay that I had developed over the years, by closely collaborating and engaging with colleagues and activists on the ground. I was keen to find ways in which my scholarship would transcend the dreaming spires of Oxford and help make a difference on the ground—which is where it really mattered in the end.

Receiving the death threats demonstrated, paradoxically, that my activist scholarship had been successful. The consequence, however, was that the dynamics of impunity that I had been analyzing for so long entangled me completely. As a recipient of death threats that Uruguayan authorities had no intention whatsoever to investigate—whether in 2017 or today—I had been drawn into the very impunity that was the object of my research. With the passage of time and the continued lack of answers, I began to experience—on a small scale—some of the consequences of the impunity that victims of the Uruguayan dictatorship had faced for decades.

Soon after the threats, on March 1, 2017, the Inter-American Commission on Human Rights condemned what was happening in Uruguay and noted the importance of prosecutions for serious crimes committed during the dictatorship in order to ensure access to justice for the victims (IACHR 2017). Two years later, with no progress on the horizon, on February 27, 2019, the Commission reaffirmed its concern about Uruguay's failure to investigate the death threats. It urged the state to ensure timely, thorough, and diligent investigations to establish and punish their perpetrators and masterminds, remarking that those threats "could increase the risk of impunity in cases linked to human rights violations in Uruguay" (IACHR 2019).

Inspired by what I had studied for years, I attempted to push back against impunity. Nine of the people who had been threatened—myself included—presented a petition to the Inter-American Commission against Uruguay in early February 2019, claiming the violation of several of our human rights, including the right to judicial protection (article 25 of the American Convention on Human Rights). We denounced the lack of progress in the investigation of the criminal case relating to the death threats we had received. We placed this in further context by showing how other human rights defenders and judicial authorities had also been threatened—most notably, the threats, break-in, and theft of equipment from the offices of the Forensic Anthropology Group, a specialized team that conducts excavations within military premises in the search for the disappeared in Uruguay, which had occurred over the 2016 Easter holiday.

While the consideration of the petition by the Commission is likely to take many years, it has already had an impact on Uruguayan authorities, who wish to maintain the country's reputation in human rights. A few months after the Commission formally notified the petition, we saw initial signs of progress. In September 2021 (well over four and half years after the threats), a 34-year-old medical student was charged with being the leader of the Comando Barneix and is currently awaiting trial. The petition is key to maintaining pressure on the authorities in Uruguay to investigate all the perpetrators and masterminds behind the threats. While the charged student might have been the person who sent the email, given his knowledge of the deep web and TOR platform, which was used to avoid leaving a footprint, he does not fit the profile of the masterminds behind these threats—both in my view and that of many of the other people threatened. Impunity is still looming over our criminal case, and we might never know who threatened us.

## Final Thoughts

If I could travel back in time to 2017 and tell my old self that the project would, eventually, be fine, I do not think that she would believe me. On many occasions, especially in the early months, dealing with the consequences of the death threats and keeping the project going seemed like an impossible task. But all the people I met during my years researching impunity in the Southern Cone have shown me what resilience is really about: to keep going even when everything seems to conspire against you.

My dedication to activist scholarship, which had put me at risk in the first place was, eventually, vindicated. Not only I did complete the project, despite significant delays, but *The Condor Trials* book was finally released in 2022 and went on to win the 2023 Juan Méndez Book Award for Human Rights in Latin America. Notably, the research I fought so hard to conduct has also had unprecedented impact, which is very close to my heart. Key findings from the database on the victims of transnational repression in South America in the 1970s were used by the Inter-American Commission and the Inter-American Court of Human Rights in 2019 and 2021 respectively in the *Julien Grisonas Family vs. Argentina* case—an emblematic Operation Condor case in which I was an expert witness for the Julien siblings, and in which Argentina was eventually found internationally responsible for the atrocities suffered by the family. In 2023, I served as an expert witness in two additional Condor-related cases. In February, I appeared before Rome's Criminal Courts, where I explained to the Italian judges the dynamics surrounding state terror and transnational repression in South America, alongside the personal stories and trajectories of the three victims and the defendant in the second Condor trial in Italy. Then in May, I appeared before Chile's Supreme Court, where I illustrated the origins of the first Italian Condor trial and described the fate of four Italian-Chilean victims whose murders had been probed in criminal proceedings which concluded in 2021.

That it became possible to present insights from my research in court in support of long-standings victim struggles against impunity reaffirmed to me the significance of activist scholarship, despite everything that had happened.

Since this experience, I am much more aware of the potential implications of my methodological and personal choices and what I would label the invisible or unplanned sources of risk. Researchers might be less aware of these given their invisible nature, but they have the potential to undermine a research project as significantly as more visible threats. Invisible threats need to be taken into careful consideration before and during a research project, not only for their professional implications but also for the researcher's wellbeing.

## References

Berger, Roni. 2015. "Now I See It, Now I Don't: Researcher's Position and Reflexivity in Qualitative Research." *Qualitative Research* 15, no. 2 (April): 219–34.

Booth, Ken. 1997. "Security and Self: Reflections of a Fallen Realist." In *Critical Security Studies: Concepts and Cases*, edited by Keith Krause and Michael C. Williams, 83-119. Minneapolis: University of Minnesota Press.

Carspecken, Lucinda. 2023. "Writing Strategies in Autoethnography and Memoir: Methodological Legacies from Three Activist-

Scholars." *Qualitative Research* https://doi.org/10.1177/14687941221138403

El País. 2017. "Se suicidó el general Pedro Barneix tras saber de su procesamiento," September 2, 2017. https://www.elpais.com.uy/informacion/se-suicido-el-general-pedro-barneix-tras-saber-de-su-procesamiento

IACHR (Inter-American Commission on Human Rights). 2017. "IACHR Condemns Death Threats in Uruguay." March 1, 2017. https://www.oas.org/en/iachr/media_center/preleases/2017/021.asp

IACHR (Inter-American Commission on Human Rights). 2019. "IACHR Concerned about Failure to Investigate Threats against Justice Operators Active in Cases of Serious Human Rights Violations during Uruguayan Dictatorship." February 27, 2017. https://www.oas.org/en/iachr/media_center/PReleases/2019/045.asp

Kohl, Ellan, and Priscilla McCutcheon. 2015. "Kitchen Table Reflexivity: Negotiating Positionality Through Everyday Talk." *Gender, Place and Culture* 22, no. 6: 747–63.

Folkes, Louise. 2023. "Moving Beyond 'Shopping List' Positionality: Using Kitchen Table Reflexivity and In/visible Tools to Develop Reflexive Qualitative Research." *Qualitative Research* 23, no. 5 (October): 1301-18.

La Porte, Jody. 2014. "Confronting a Crisis of Research Design." *PS: Political Science & Politics* 47, no. 2 (April): 414-17.

Lessa, Francesca. 2022. *The Condor Trials: Transnational Repression and Human Rights in South America*. New Haven: Yale University Press.

Mac Ginty, Roger, Roddy Brett, and Birte Vogel, eds. 2021. *The Companion to Peace and Conflict Fieldwork*. London: Palgrave Macmillan.

Mendez, Melissa. 2023. "Partisanship and Positionality in Qualitative Research: Exploring the Influences of the Researcher's Experiences of Serious Crime on the Research Process." *Qualitative Research* 23, no. 1 (February): 92–107.

Nordstrom, Carolyn, and Antonius Robben, eds. 1996. *Fieldwork Under Fire: Contemporary Studies of Violence and Culture*. Los Angeles: University of California Press.

Sriram, Chandra Lekha, John C. King, Julie A. Mertus, Olga Martin-Ortega, and Johanna Herman, eds. 2009. *Surviving Field Research: Working in Violent and Difficult Situations*. Abingdon: Routledge.

Schulz, Philipp. 2021. "Recognizing Research Participants' Fluid Positionalities in (Post-)Conflict Zones." *Qualitative Research* 21, no. 4 (August): 550–67.

Valenzuela-Fuentes, Katia. 2019. "Militant Ethnography and Autonomous Politics in Latin America." *Qualitative Research* 19, no. 6 (December): 718–34.

QMMR Qualitative & Multi-Method Research

Notes from the Classroom

Qualitative and Multi-Method Research | Fall 2022, Vol. 20.2 | https://doi.org/10.5281/zenodo.8326533

# Reimagining Research Design Instruction: Student and Teacher Reflections on the Reverse Research Design

Phillip M. Ayoub
*University College London*

Jaya Duckworth
*Occidental College*

This piece is a follow-up on a pedagogical exercise called the "reverse research design" (Ayoub 2022).[1] As a teaching tool, the reverse research design involves students stepping into the shoes of a published author and transporting themselves back in time to craft a grant proposal for an already-concluded study. This hands-on exercise guides them through the intricacies of research design while temporarily easing the anxiety of formulating their own research question and project. At the request of the QMMR editors, we

1 We summarize the assignment and parts of the argument based on the Ayoub (2022) study, which we recommend being read in conjunction with this follow-up piece.

restate the goals of the original exercise but also build upon it by developing additional strategies for productive use in the classroom. To that end, it is vital to incorporate student viewpoints and feedback, which we do below by uniting the perspective of a student (Duckworth) and an instructor (Ayoub). Our endeavor is thus to offer a more personal reflection about our experiences with this activity, coupled with a discussion of wider student feedback.

We begin first by summarizing the nuts and bolts of the reverse research design assignment, followed by a reflection from the classroom on how to implement it effectively. This section incorporates a student and instructor perspective. Finally, we analyze student evaluation feedback to offer some descriptive evidence of the assignment's utility in sharpening students' analytical skills for social science research.

## Summary of the "Reverse Research Design"

In brief, the reverse research design serves as a bridge, effectively combining two key expectations placed on students within the academic setting: first, the requirement to read and critically engage with the works of established scholars, and second, the need to develop their own research skills. While we often teach these expectations separately, the reverse research design directly deploys one expectation to build on the other. This section explains the concept of reverse research design and provides a brief summary of the teaching resources and methodologies employed for its implementation.

The reverse research design involves students retracing the research process of a published author by imagining themselves as the author writing a grant proposal for the study. A key element of this exercise, also discussed from the student perspective below, is that it allows students to work through the steps of research design without the pressure of formulating their own research question and project. The teaching tool includes three steps to teaching research design. Step 1 begins with the instructor introducing the core components and purposes of research design. Step 2 is the reverse research design assignment itself, where students create their own research proposal based on an existing work. Step 3 involves an original research design assignment, where students formulate their own research question and proposal.

The crucial bridging step here is Step 2, where students work with a book or article they have become familiar with in the course (from the course outline), identifying the strategies the author employed to create a finished work.[2] Familiarity with the book or article reduces students' anxiety about designing their own research projects and allows them to creatively explore the author's profile and experiences during the research process. In the assignment, students envision themselves as the author at the initial research design stage. They write a grant proposal to a foundation seeking funding for the research that led to the published book or article. The proposal should include a research design outlining the project in the student's own words, with some creative freedom allowed, as they do not know all elements of the author's process. The components of the research design they are asked to identify and address include the puzzle, research question, and argument; data collection and methodology; feasibility (here students also research the author's language skills or methods training from their CVs); and significance of the project. Ideally, most of these elements (especially the first three) should be explicit in the original article or book. This exercise becomes useful for understanding the writing process, comprehending the assigned readings by carefully dissecting them, and preparing for future research projects of their own. To that end, the exercise can also be assigned in the form of a Fulbright Grant proposal if the instructor wishes. This helps students imagine themselves applying for actual grants, simultaneously demystifying that process and increasing the odds that they will submit such grant applications of their own (see below). A full breakdown of the steps and an example of the assignment handout can be found in Ayoub (2022).

The task can also be implemented individually or as a group assignment. While there is some benefit to grading if students work on the assignment on their own at home (as it demonstrates their degree of engagement with the course material and places them in the driver's seat at the outset) it has also worked well as a group assignment. In some years, Ayoub assigned it in teams of four students, where they started the project in class and worked on it for a week outside of class. This can be more manageable for instructors who have a packed syllabus already and lack the time to incorporate it as a separate individual assignment, or for instructors in large courses where the grading lift of individual assignments would be too high. In the group version, teams produce the grant proposal together, which lowers the output of the additional grading material when working this assignment into a syllabus. For Duckworth, working in groups was very effective. She found it helpful to talk through the elements of the research design with peers, dissecting the project together. Of course, one downside

2 Usually, all students work with the same text (one the instructor has vetted to ensure it has a clearly explained methodological approach) so they can compare their research design after. That said, some instructors may allow students to select any piece from the course outline.

of the group setting may be that it diminishes the effect of fully embodying the author oneself and, as with any group task, there remains a risk of a few students dominating—creating a setting where marginalized students remain marginalized.

In sum, we hope the tool addresses some of the challenges of teaching research design in undergraduate social science courses by proposing a pedagogical tool to bridge the gap between reading established scholars' work and students conducting their own research. Importantly, this small assignment also helps address the importance of teaching research design, despite the challenges of doing so. As we build upon below, instructors—who typically have been working with the research process for many years —often forget the felt intimidation that "doing one's own research" provokes in a student. Also, the pressure of coming up with one's own exciting and "uncharted" question is a real handicap to beginning a project. By having those elements provided by a published author (i.e., removing some of the mental impediments to getting going) students can work through the process initially. Later, having completed reverse research design on paper, they feel much more confident in wearing the hat of the researcher themselves.

## Why and How It Matters: A Student Reflection on Implementation

The process of moving from intimidation to confidence while undergoing this task is described by many students, including one of the authors here. To be sure, a primary goal of this assignment is to demystify the daunting process of research design. While undergraduate institutions advertise their numerous opportunities for student research, many students of the social sciences—particularly students who are historically marginalized—view the research process as unfamiliar, intimidating, or exclusive to science, technology, engineering, and math (STEM) fields (Ayoub and Rose 2016). This was certainly the case for Duckworth, who was completely unfamiliar with the research process when she entered Occidental College as an undergraduate. At small liberal arts colleges, where research methods courses are not always required in social science departments, students may advance into upper-level courses without ever being introduced to the process of research. This is also true of many departmental curricula in many larger universities. Duckworth herself experienced this at Occidental, where her first course dealing with research design or method was Ayoub's Comparative Social Movements seminar, which she took in her third year. Indeed, Ayoub's use of the reverse research design in many topical courses (like the one on social movements) was informed by the fact that there was little opportunity at the college for students to hone their own research skills before dropping them into the deep end of the pool during their senior comprehensive project at the end of their senior year. Many students felt more needed to be done, and the department had few resources and no tradition of teaching methods until recently (a 2-credit methods course is now offered, though not required, as part of the major curriculum).

The reverse research design assignment thus offers an entry point for students in Duckworth's position who are new to research, lack confidence in their ability to conduct research, and are unsure of where to start. The exercise feels manageable and requires little external preparation, as students are already familiar with the material they use for the assignment and can simply work backward to break down its methodological components. This eliminates the especially daunting task of identifying a puzzle and generating one's own original research question, which Duckworth and many of her peers found to be the biggest challenge as undergraduate researchers. The pressure of devising a research question that "had never been studied before" to solve a puzzle that "had never been answered before" felt paralyzing (and somewhat ridiculous) to Duckworth. Even though her instructor insisted that research questions are rarely fully untapped and expressed skepticism about the hunt by undergraduates for true "gaps in the literature," students still felt that pressure. Duckworth would ask herself: How could I possibly come up with something new that no researcher has ever thought of before? What makes me qualified to contribute to a conversation between experts who have dedicated their careers to this topic? The reverse research design assuages these concerns—or at least defers them until later in the process, when students feel more confident in their skills and have a more realistic understanding of what is expected of undergraduate researchers (e.g., by charting the years-long process of published work).

After completing the reverse research design assignment, Duckworth felt more empowered to begin an original research design than she did in other research-based undergraduate courses. In fact, she later took a similar upper-level research course that did not use this intermediary assignment during the research process, and she noticed herself and her fellow students struggling considerably more. Many students in that course were able to identify a topic they were passionate about but found it difficult to translate it into a specific and original research question, defaulting instead to vague and exploratory questions. Further, without reading course materials with a specific attention to methodology, students had a weaker grasp of how to craft a research question and puzzle and what the different elements of a research design were. Their doubts about their own ability to contribute something meaningful to the literature remained paramount. Duckworth herself spent many hours sifting through literature that had not been assigned in class in order to develop an original research question and puzzle, creating significant extra work and anxiety about the research process. The reverse research design mitigates many of these roadblocks by utilizing familiar literature, teaching students to analyze texts specifically for their methodologies, and breaking down the research process into digestible segments that students can then apply to their own work.

## Seeing Oneself as a Researcher

By following an author's biography (in the social movements course, some students write about the author's experiences in activism, or discover that they were first-generation students, etc.) the roadblocks start to chip away for many students. To that end, because a primary goal of this assignment is to demystify the research design process for students who may feel alien to research, we especially recommend using work written by scholars who are themselves underrepresented in the field. Work by scholars who address their status as first-generation scholars, or scholars of color, or scholars marginalized in fields where their abilities, gender, or sexuality are underrepresented can be empowering. In the year of Duckworth's class, we worked with a book by Chris Zepeda Millán (2017) called *Latino Mass Mobilization*. The book was directly related to the course material, dealing with many of the theories of social movements we had been studying, but it was also useful for the discussion of Zepeda Millán's positionality. In the appendix, he discusses his experiences as a first-generation scholar of color in political science at great length, describing both the challenges and hidden benefits of that role. In particular, Zepeda Millán discusses how his own Latino identity and his connection to immigrant communities should be viewed not as a conflict of interest or a threat to "objectivity," but instead as an asset that grants him privileged access to and credibility among the activists and organizers he interviewed for his book. Further, he argues that his personal investment in the success of the immigrant rights movement motivates him to be a more thorough and accurate researcher, in order to generate meaningful results that can be useful to both scholars and the communities he studies. Given that our course took place in Los Angeles, this element of the exercise had a powerful resonance within debates in which many students were engaged in their everyday lives and conversations.

For Duckworth, this perspective helped address some of the impostor syndrome she felt as a woman of color in higher education. As a scholar-activist herself who was highly involved in movements for racial and gender justice, she appreciated the idea that research need not be apolitical, realizing that her identity and activist background could actually strengthen her research capabilities. Thus, Zepeda-Millán's approach not only carves out space for marginalized scholars to contribute to a white and male-dominated domain, it also illuminates how one's marginalized identity and existing set of experiences can actually be powerful assets in the research process. This perspective increased students' confidence and enhanced their methodological framework by combining research design with feminist and queer methods. In the assignment, because they pretended to be the author, many students— including first-generation and marginalized students—were able to identify with Zepeda-Millán and see themselves in his research, which made playing the role of a "scholar" more transformative. Ultimately, when this assignment uses the work of marginalized scholars, it functions not only as a pedagogical tool but also as a vehicle to foster greater belonging for marginalized students in academia.[3]

## Overview of Student Feedback

This final section provides some descriptive evidence of the assignment's effectiveness in improving students' understanding of research methodology and independent inquiry. Table 1 shows the mean scores of the four relevant questions from student evaluations in two courses that used the reverse research design assignment at Occidental College.[4] While the two courses make a small sample of descriptive data from which we should interpret cautiously, they do suggest that the assignment resonated well with students. In both courses, a majority of students either agreed or strongly agreed that the course improved their ability to analyze and synthesize information, write clearly and effectively, read critically and analytically, and think critically and

3 Of course, this will depend on the topics of the course, but other books and articles by marginalized scholars I have used include those by Tina Fetner (2008), Brian Harrison and Melissa Michelson (2017), Rupp and Taylor (2003) and Phillip Ayoub (2016, 2014).

4 The pattern was largely the same when it was implemented at Drexel University and Cornell University. We do not include those evaluations here because the survey questions were different. This assignment was also taught a third time at Occidental College in spring 2020, where quantitative evaluations were canceled by the college due to the COVID-19 pandemic. Finally, we also compared the data on these four analytical questions to other courses Ayoub taught without assigning the 'reverse research design' at Occidental College. Holding constant the instructor (but not the course), we notice a systematic difference in the question of student self-assessment of their "analytical skills" between courses that used the 'reverse research design' assignment and those that did not.

analytically. Across all four questions, the mean score was above 6.00, on a scale from 1 (strongly disagree) to 7 (strongly agree), and consistently above the college and departmental averages.[5] Such outcomes were similarly reflected in students' qualitative comments, where many noted improvements in their writing skills and highlighted the utility of the course in preparing them for future research. One student wrote, "I learned how to effectively write a research design, synthesizing previous literature and theory and effectively building my own research proposal. This greatly strengthened my ability to write concisely and clearly with the goal of proposing a research project, which I can use for future opportunities." Several students similarly noted how the methodological training from the course had equipped them well to begin their senior comprehensive research project, and some wrote that they intended to use their research design from the course to apply for a Fulbright. "I am planning to use my work in this class to apply for a Fulbright," wrote one student. Another said, "This course taught me how to write a research design and create a Fulbright application. These are two skills I didn't previously have, but I'm so thankful that I have them now!"

Along with highlighting improved writing and research abilities, several students commented on the course's gradual pacing of the research process, which eased their stress and made their own research design feel more manageable. One student wrote, "Because we spent more time truly understanding what social movement research papers are supposed to look like (through Zepeda-Millan and Prof. Ayoub's own book[s]), writing the final paper was much easier than I had originally anticipated." Another reflected, "the pacing of the research proposal was very effective, [and] I really appreciated how we gradually edited our work throughout the semester. It made the assignment much more manageable and strengthened my writing." Thus, students appreciated intermediary assignments like the reverse research design, which gave them ample time to deconstruct the research process and empowered them to begin their own project. Many also noted how their own research plans for the major had been strengthened. For example, one said they "feel more than prepared to take on the senior seminar and my final year in the major," and another noted that they now had the "tools to launch [their] senior comprehensive projects, which [they are] very grateful for."

Table 1: Summary of Student Evaluations
*Note: Mean scores are reported, Scale from 1 (strongly disagree) to 7 (strongly agree)*

| | Spring 2019 | Fall 2021 |
|---|---|---|
| This course improved my ability to analyze, synthesize, and/or apply information regarding the subject | 6.86 | 6.86 |
| This course increased my knowledge, skills, and/or personal development in the following areas: **writing clearly and effectively** | 6.75 | 6.46 |
| This course increased my knowledge, skills, and/or personal development in the following areas: **reading critically and analytically** | 6.88 | 6.07 |
| This course increased my knowledge, skills, and/or personal development in the following areas: **thinking critically and analytically** | 7.00 | 6.29 |

Finally, several student evaluations mentioned how the course had inspired passion and confidence in their abilities as researchers. One student wrote that the course "inspired [them] to try and consider doing research in the future," and another reflected, "I have always been insecure about my writing, but the amount of time we have had in class to discuss our [draft] papers has been tremendously helpful for myself and my peers."[6] Several highlighted its practical utility, calling it "super practical" in a way that "will help me well beyond the class setting,

5 We opted to report the mean scores. Median scores were consistently higher.

6 To circle back to the above, the reverse research design is a preliminary task that kickstarts the writing process for students' own research design papers, referenced in a student's quote here. After the reverse research design assignment is complete, students then work incrementally to draft each section of their own research design (i.e., research question, methodology, feasibility, etc.) throughout the remainder of the course, which they eventually submit as the final paper for the course. Over the course of several weeks, they participate in periodic peer review sessions during class, receiving feedback on each new section of their research design. This intentionally gradual process, which begins with the reverse research design, has been appreciated by many students.

leaving me with a lot of good skills I can later use in my college career and beyond." One commented, "I believe the content of this course will significantly impact my future career and involvement." Students thus not only improved their analytical abilities, writing, reading, and critical thinking skills, they also began to view themselves as more capable researchers, increasing both their confidence and excitement about participating in research in the future.

## Conclusion

Our hope is that this joint student and instructor reflection offers fresh ideas for how to implement the reverse research design. We have presented a summary of the assignment as a complement to the Ayoub (2022) article, followed by a first-hand reflection on our experiences both implementing it and undertaking it in the classroom. As a pedagogical approach, we find the reverse research design to accomplish multiple learning outcomes, including facilitating the understanding of course material, introducing the logic underlying research design, fostering adeptness in reading and comprehending social science literature, and inspiring independent inquiry among students. Finally, we have offered some anecdotal descriptive evidence from students about how the assignment teaches them both a topical area and research design methods, while building their confidence as independent researchers. Students consistently laud it as fundamental preparation for their subsequent comprehensive thesis writing, grant applications, and comprehension of social science literature. In sum, our experience with the assignment—and that of colleagues that have implemented it—suggests it is a simple but effective tool and its implementation has yielded successes. Despite its apparent simplicity, the exercise confers benefits to students, enhancing their ability to conduct research meaningfully. By equipping students with the skills necessary to formulate and address novel questions, we can contribute to their academic growth and potential.

## References


Ayoub, Phillip M. 2022. "Reverse Research Design: Research Design in the Undergraduate Classroom." *PS: Political Science & Politics* 55, no. 2 (April): 424–28.

———. 2016. *When States Come Out: Europe's Sexual Minorities and the Politics of Visibility*. Cambridge: Cambridge University Press.

———. 2014. "With Arms Wide Shut: Threat Perception, Norm Reception, and Mobilized Resistance to LGBT Rights." *Journal of Human Rights* 13 (3): 337–62.

Ayoub, Phillip M., and Deondra E. Rose. 2016. "The Scholarly Importance of Studying Issues

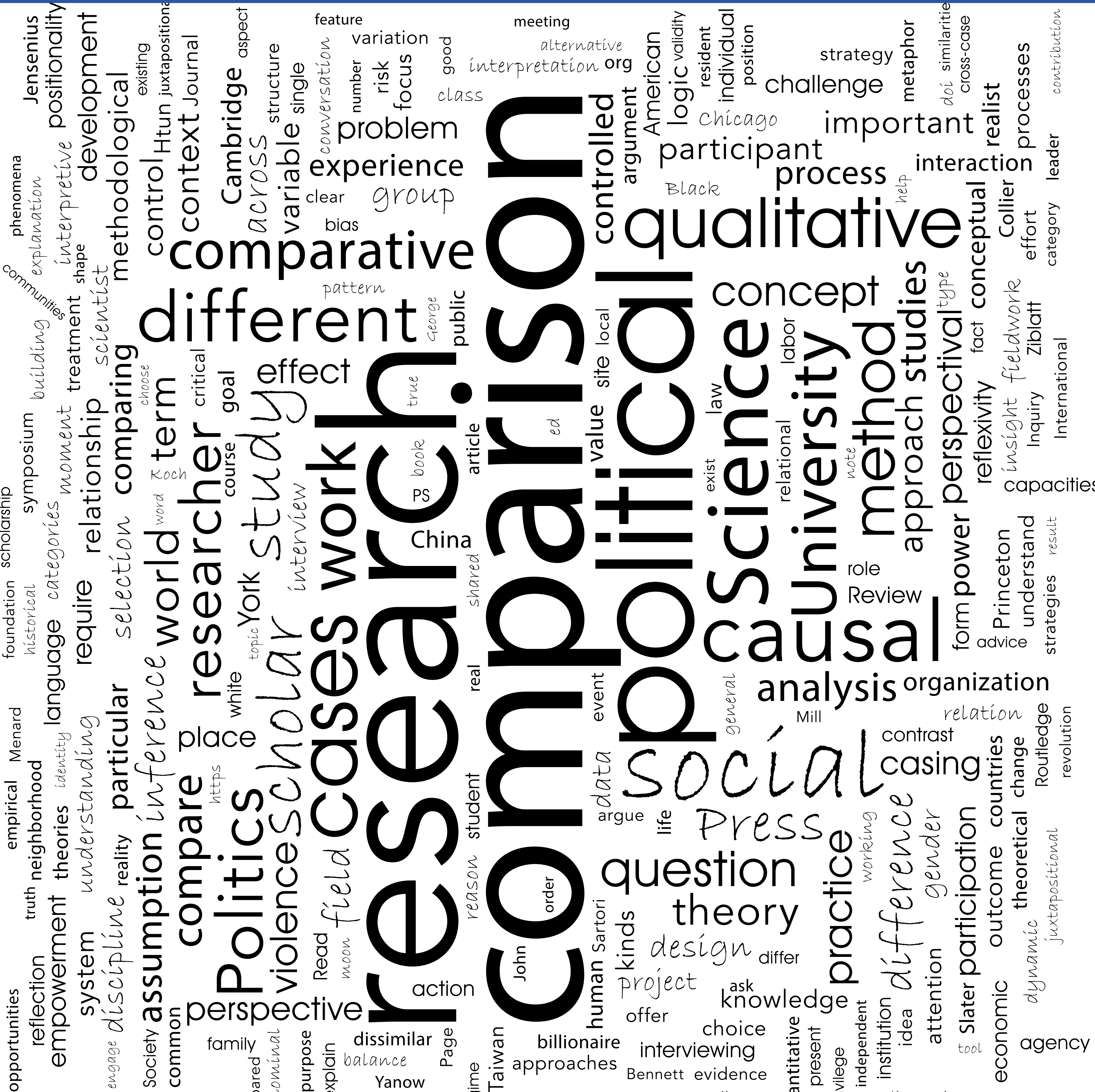
Related to One's Own Identity." *Inside Higher Ed*, April 14, 2016. www.insidehighered.com/views/2016/04/14/scholarly-importance-studying-issues-related-ones-own-identity-essay.

Fetner, Tina. 2008. *How the Religious Right Shaped Lesbian and Gay Activism*. Minneapolis: University of Minnesota Press.

Harrison, Brian, and Melissa Michelson. 2017. *Listen, We Need to Talk: How to Change Attitudes about LGBT Rights*. Oxford: Oxford University Press.

Rupp, Leila J., and Verta Taylor. 2003. *Drag Queens at the 801 Cabaret*. Chicago: University of Chicago Press.

Zepeda-Millán, Chris. 2017. *Latino Mass Mobilization: Immigration, Racialization, and Activism*. New York: Cambridge University Press.




QMMR
Qualitative &
Multi-Method
Research