# QMMR

## Qualitative and Multi-Method Research

### Fall 2022 | Volume 20.2

# QMMR
Qualitative &
Multi-Method
Research

# Fall 2022 | Volume 20.2
# Table of Contents

# QMMR
Qualitative &
Multi-Method
Research

# Letter from the Editors

We couldn't be more excited to see this issue out in the world. This is our first one as solo editors of QMMR. We would like to pay tribute to our predecessor, Jennifer Cyr, who did a terrific job steering the ship and helping us with the transition. Our goal is to build on her work, making QMMR a diverse venue where junior and senior scholars from all corners of the globe come together to share cutting-edge methodological developments as well as their experiences as teachers and practitioners of qualitative and mixed methods research. We envision a publication that is radically practitioner-orientated at its core, one that makes epistemological and methodological discussions accessible, appealing, and useful for the thousands of scholars who rely on qualitative and mixed methods.

The contributions to the current issue are in line with this goal. In the pages below, readers will find three original articles. First, Daniel Solomon, a Ph.D. candidate at Georgetown, explains the meaning and significance of "paradigmatic cases." In his view, researchers should employ this terminology only when their cases undermine – not affirm – the core assumptions and empirical research standards of social science programs.

Second, in an effort to decolonize research design, Mneesha Gellman proposes an innovative collaborative methodology for engaging stakeholders. Specifically, she draws on years of experience conducting fieldwork among indigenous peoples in Northern California and Oaxaca to outline the contours of this approach and alert us of its challenges.

Third, Anthony DeMattee, Nick Gertler, Takumi Shibaike, and Elizabeth A. Bloodgood provide excellent advice for scholars who rely on legal texts such as statutes and regulatory rules as primary sources. They argue that those working with documents in multiple languages may occasionally get "lost in translation." This is because available strategies to translate legal texts are either cost-prohibitive or error-prone. Based on available software, the article evaluates several solutions to the "laws-in-translation problem."

The issue also features a timely symposium on fieldwork in post-pandemic times. It reflects on the ethical and practical lessons learned by a group of researchers who conducted qualitative research during Covid. The essays discuss the extent to which the adaptations most of us had to make during that period should continue in our new era. In other words, what did we learn about fieldwork safety during an unusually uncertain and difficult period that may also apply to fieldwork more generally? This is an important question that underscores the fact that issues of fieldwork safety and ethics should never take the backseat, even when we collectively lower the guard. This symposium is a great follow-up to the contributors' recently published volume, *Safer Field Research in the Social Sciences: A Guide to Human and Digital Security in Hostile Environments* (SAGE, 2020).

As part of our effort to draw practitioners and teachers into the QMMR community, and thus serve a broader audience, we inaugurate two new sections. In addition to original articles and multi-author symposia, from now on each issue will include *Notes from the Field* and *Notes from the Classroom*. In *Notes from the Field*, we want to hear from graduate students and established researchers who are currently in the field or have recently completed rounds of fieldwork. The essays will consist of vivid reflections

about the personal, logistical, and methodological challenges encountered during the research process, or discuss the use of data-gathering and organizing techniques in specific settings. In sharing their stories, contributors to this new section will address important ethical and methodological questions, as well as offer practical advice.

The inaugural *Note from the Field* by Will Freeman tells the riveting story of a researcher grappling with questions of access, meaning, and validity during a Ph.D. dissertation field trip to Peru. There he interviewed elites involved in, and affected by, a historic wave of corruption prosecutions. Will reflects on the challenges posed by interviewees' metanarratives and how he struggled to read between the lines.

*Notes from the Classroom* is a space for scholars who teach qualitative and mixed-methods courses to share their experiences and general wisdom with our readership. We will publish advice on how to (and how not to) teach certain topics or techniques, including examples of innovative syllabus design, in-class exercises, and/or formative/ summative assessments.

The inaugural *Notes from the Classroom* by Julia Lynch is a perfect example of what we envisioned. Here we have an incredibly experienced instructor walking us through an inspiring method to teach concepts and concept formation at the graduate level. Julia distills the advantages of active learning exercises: they facilitate the transition from conceptual to practical understanding, deliver immediate impact, and level the playing field between students with different levels of prior knowledge or experience.

Before we let readers dig into this rich material, we would like to encourage the QMMR community to continue submitting 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 revamped website: https://www.qmmrpublication.com

We look forward to walking this journey with you over the coming years!

Ezequiel Gonzalez-Ocantos          Juan Masullo J.
*University of Oxford*                    *Leiden Unviersity*

# Finding Our Finches: Paradigmatic-Case Research in Political Science

Daniel Solomon
*Georgetown University*[1]

## Introduction

The revolutions that swept across the Middle East and North Africa during the first half of 2011 rattled both political science research and policy consensus about political and social development in the region. For many political science researchers, the so-called Arab Spring upended common assumptions about the conditions necessary for both authoritarian durability and large-scale social upheaval, and the characteristics of regimes that illuminate changes in these phenomena (Bellin 2012). For policy practitioners, the revolutions and their broader consequences upset prevailing perceptions about political instability and repression in the region (Laipson 2011).

These sorts of standard-shifting cases underpin the practice and mythology of scientific research across disciplines. Consider two well-tread accounts of scientific discovery beyond political science: John Snow's (1855) London water pumps and Charles Darwin's (1859) finches. Both provided foundational models of causal relationships in their respective disciplines: the former for the spread of cholera and other infectious diseases; the latter for the process of evolutionary adaptation. These initial inquiries also transformed the empirical study of their respective research topics. Snow's otherwise-similar London neighborhoods challenged conventional assumptions about the invariable link between disease and impoverishment; meanwhile, Darwin's Galapagos finches showed how environmental conditions shape organisms' expressive traits. Since Snow's and Darwin's

initial work, researchers have returned to the two cases to revise core concepts in research about public health and evolutionary ecology, respectively (Koch and Denike 2009; Grant and Grant 2002). Historical accounts of the scientific communities that contributed to the two researchers' empirical work also demonstrate that these presumed turning points instead reflect continuity in the scientific study of epidemiology and ecology (Worboys 2000; Berry 2014).

Versions of the water pump and the finch abound in contemporary political science. In many of our various research programs, political scientists rely on a limited universe of exceptional examples that stand in for a well-established consensus about the phenomena that we study. In some circumstances, these cases can be a form of professional jargon. A scholar of diplomacy, for example, expects that both their research colleagues and their policy counterparts will recognize "appeasement at Munich" as a specific pathway of strategic restraint against ascendant hegemons (Beck 1989). But these cases are not only academic shorthand; researchers also attribute analytic weight to their findings. Authors of hundreds of studies published in top political science journals over the last half-century have justified their subjects of inquiry as *paradigmatic*.[2] In general, researchers use the language of "paradigmatic cases" to imply that conclusions from a case or cases under study warrant greater attention than other instances of the same phenomenon or causal process.

Despite these recurring references, the political science discipline lacks clear guidelines for paradigmatic-

case research as a mode of empirical analysis.[3] Although the discipline has a large vocabulary of terms that signal a case of outsized importance, few match up to the worldview-collapsing standard that paradigmatic-case research implies. For some researchers, paradigmatic cases speak to conceptual models; for others, to theoretical relationships between variables; and for a third group, to conventional assumptions among policymakers and other practitioners.[4]

These terms are easy to confuse with other ways of referring to "very important" examples, such as the cases most affirmative of a theorized relationship that Eckstein (1975) describes as "crucial." Eckstein's own crucial case—Norway's stable, socially cohesive democracy—is an example in point. For Eckstein, Norway is crucial because it most coherently aligns with the anticipated positive relationship between social cohesion and democracy (Eckstein 1975; 1966). Eckstein does not demonstrate, however, that Norway has any meaningful influence over the broader research agenda on democracy or social cohesion. A paradigmatic case study would also show that the Norwegian model has influenced the core concepts and empirical study of these topics. Eckstein's maximally-congruent Norway may be a crucial or important example, but the scientific consequences that Eckstein demonstrates fall short of paradigmatic. Unlike crucial cases, only paradigmatic cases imply that the case study's conclusions merit the wholesale transformation of the entire scientific worldview or consensus that underpins a research program.

In this essay, I examine how empirical case studies may contribute to research paradigms or programs in political science. I proceed in three parts. First, I elaborate on a practical definition of paradigmatic cases and ground this definition in a brief summary of philosophy-of-science debates about stasis and change in communities of scientific knowledge. Second, I discuss two necessary standards of evidence—one logical, the other sociological—that researchers should use to demonstrate that a case is "paradigmatic." Third, I conclude by briefly discussing how a common vocabulary of paradigmatic-case research contributes to research about social science programs. Defining the analytic boundary between paradigmatic cases and other very-important types is not simply a matter of terminological minutiae. The task of identifying paradigmatic cases and

how they contribute to our research programs advances the collective development of scientific knowledge about politics by making explicit one source of foundational disagreement within empirical research programs that researchers often fail to specify.

## The Paradigmatic Case and its Implications

Paradigmatic cases are the observable markers along the "route to normal science" (Kuhn [1962] 1970). In this essay, I define paradigmatic cases as *discrete instances of an empirical phenomenon that may undermine the validity of research programs*. Before defining the relationship between paradigmatic cases and research programs, I clarify two main components of this definition: namely, "discrete instances of an empirical phenomenon" and "the validity of research programs." The first refers to cases that are characterized by (1) boundedness and (2) external comparability to other instances beyond the case (Gerring 2004). A case may be well-bounded when it occurs at a specific time or in a specific location that separates it from other cases. In this context, a case's boundedness may also refer to other forms of classification behind historical time or geography. For example, Hathaway (2007, 589) refers to human rights treaties as a "paradigmatic hard case" of state commitments in international relations. Here, the "discrete instance" in question is a type of state behavior—treaty adoption—rather than a specific historical episode.

Cases are externally comparable because they are all instances *of something*. The empirical phenomenon that the researcher examines and the classification system that they use to study it determines the specific bounds of the case (Gerring 2016). These classification systems influence the attributes of the case that warrant empirical study. In a study of democratization in nation-states, the salient empirical characteristics of, say, contemporary Nigeria are those—its regime, its historical political development, or its social movements—that make Nigeria comparable to other nation-states. These characteristics differ from those that a researcher might examine in a comparative study of informal labor markets, or a study of imperial collapse in which contemporary Nigeria would not feature. Studies of disaggregated *units of observation* within a case—for example, municipalities within a single country—can demonstrate how the political conditions of the macro-level case shape

---

3 Other efforts to define and establish criteria around paradigmatic cases beyond the political science discipline bear noting. In particular, see Flyvbjerg (2006), who dismisses the possibility of a consistent, universal standard of paradigmatic-case selection; and Mills, Durepos, and Wiebe (2010), whose definition resembles the language of "crucial case" (Eckstein 1975) research in political science.

4 More specifically, these different phrases imply that cases are paradigmatic because they: (1) are typical examples of an empirical phenomenon (e.g., Croke et al. 2016); (2) demonstrate the "on the line" value of a statistical relationship between two variables (e.g., Sa'adah 2006); (3) illustrate an exception to a theoretical rule (e.g., Hathaway 2007); (4) provide a benchmark or model for policy practice(e.g., Balcells 2010); or (5) should or do demand the attention of policy practitioners or the general public (e.g., Manekin 2013) .

within-case units, in addition to the meso- or micro-level determinants of cross-unit patterns (Giraudy, Moncada, and Snyder 2019). In these circumstances, however, the case encompasses these disaggregated units.

To clarify the second component of paradigmatic cases—their relationship to the "validity of research programs"—I adapt Lakatos's ([1978] 1980) definition of research programs to refer to a series of empirical inquiries about an observable phenomenon based on widely-shared, inviolable assumptions about that phenomenon's general properties.[5] Researchers working towards a common program generally agree on three main characteristics of their collective inquiry. First, they agree about the necessary conditions that constitute their research subject. These necessary conditions broadly align with Goertz's (2006, 5) approach to "concepts" as "theories about the fundamental constitutive elements of a phenomenon." Without a common concept, researchers have no common understanding of what general category of phenomena structures their study. Second, researchers should reasonably expect that their fellow researchers may agree about how those necessary conditions apply to the "real world"—that is, they should share a common standard of observation. Without a common standard of observation, the universe of cases that comprise the research program's empirical inquiry is an ever-shifting target. Lastly, researchers agree about the levels of analysis at which the program applies that observational standard.

The basic conceptual framework of scientific research programs and the resulting Lakatosian "methodology" for evaluating their progress does have some purchase in the social sciences. The robust contemporary research program surrounding the social construction of identity is one such example (Chandra 2012). The concepts, observational standards, and levels of analysis that social constructivists apply to the study of identity are unrecognizable to and incommensurate with earlier essentialist theories that associate social categories with immutable biological traits. This transformation means that contemporary social-science research relegates essentialism—the dominant mode of inquiry into human social relationships for much of the 19th and early 20th centuries—to the scientific dustbin. The concepts and measurement standards that essentialist theorists of race, ethnicity, or gender employ have no role in testing nor explaining how social relationships lead identity categories to emerge, change, or influence other forms of social and political behavior.

What political scientists describe as "paradigms" more often refers to narrower theories or subjects of policy consensus. The oft-mislabeled international relations "paradigms," for example, are not full-fledged research programs because they do not rest on mutually exclusive assumptions about the behavior of states and other actors in the international system (Jackson and Nexon 2009). Research about the relationship between economic development and regime type illustrates this distinction. This program rests on the central assumption that the economic circumstances of regimes shape their patterns of survival and change. The most common—and commonly disputed—gloss on the association between development and political change, modernization theory, is simply that: one among a series of theories that encircle the inviolable assumptions of the program's "hard core." Evidence that refutes or places caveats on the *positive* association between economic development and democracy only invalidates modernization theory (Treisman 2020); it does not imply the broader program's irreconcilable opposite, that economic conditions play no role in a regime's political fortunes.

Philosophy-of-science researchers advance two competing models of the potential relationship between paradigmatic cases and research programs: one additive, the other subtractive. In Kuhn's ([1962] 1970) additive model, empirical researchers create paradigms over time by examining cases that affirm the research consensus or challenge its assumptions and implications. For Kuhn, paradigm-affirming research consists of three core features: (1) core axioms that researchers within the paradigm generally accept as true; (2) common measurement tools that enable tests of widely-accepted paradigmatic ideas; and (3) second-order extensions of the paradigm that reinforce its fundamental conclusions (Kuhn [1962] 1970, 25).[6]

Two of Kuhn's major interlocutors, Popper and Lakatos, offer separate subtractive models of research about specific cases and their implications for scientific consensus. Popper ([1934] 2002, 277) describes a process of empirical falsification in which a continuous series of "crucial experiments…[are] designed to bring about a decision between two competing theories by refuting (at least) one of them." Because this process implies incremental, decentralized tests of prevailing theories, Popper's universe of potential paradigmatic cases is

---

5 Lakatos ([1978] 1980) conceives of research programs as a combination of a "negative heuristic," a set of theoretical first-principles that comprise the program's inviolable core; and a "positive heuristic," or series of second-order implications that researchers subject to and refine with empirical tests. Although philosophy-of-science discussions sometimes use "research programme" to distinguish the Lakatosian term-of-art from the vernacular phrase, I use "research program" for clarity.

6 As Walker (2010) observes, Kuhn's "paradigms" and Lakatos's "research programs" are functionally synonymous.

theoretically infinite. Each Popperian case is equally likely to falsify a program's core ideas and findings. For his part, Lakatos refutes altogether the possibility of identifying paradigmatic cases in real time. For Lakatos ([1978] 1980, 111), the paradigmatic status or theoretical importance of any empirical inquiry is a post-hoc figment of the research program's dominance.

## What Makes a Case Paradigmatic?

Researchers may describe their case under study as paradigmatic if it meets two criteria—one logical, the other sociological. Lakatos's negative heuristic—the basic properties that undergird the research program—provides the first, logical standard with which to evaluate the program-level implications of a paradigmatic case. This standard addresses the *coherence* of the research program's central concepts. If research about a paradigmatic case can demonstrate that the three core properties of a research program are less coherent than the prevailing consensus presumes, the subsequent architecture of the research program collapses: neither its theoretical implications nor its core measures or methods of analysis may hold. The relative stability of research programs means that this logical standard amounts to an analytic "hoop test" (Van Evera 1997; George and Bennett 2005) that may only undermine—but not confirm—a research program's core ideas or their implications.

In addition to this logical standard, a paradigmatic case should also be *sociologically decisive* for the research community that steers the program. As Kuhn ([1962] 1970) observes, the final stage in the consolidation of a research paradigm is a product of actions and norms that scientific fields use to achieve consensus and reinforce disciplinary boundaries. In addressing this sociological process of consensus-making, the paradigmatic-case evidence should demonstrate that the case is the empirical keystone of the research consensus that the new inquiry seeks to overturn. To recall the introductory examples from the natural sciences, there is nothing unique about Snow's (1855) water pumps nor Darwin's (1859) finches other than their overwhelming influence on the intellectual history of their respective fields. The epidemiological process that Snow (1855) associated with the London water pumps would have been typical of any major city at the time; for Darwin's (1859) part, the Galapagos Islands were unusual but not altogether unprecedented in the biodiversity of their finch population. Looking at these cases in a new light can reveal that the original, paradigm-defining observations are attributable to incoherent concepts or a mismatch between concepts and the standards and levels of analysis that researchers use to observe them.

In doing so, they demonstrate that the conventional wisdom that the research program associates with the specific case should reflect some alternative property of the same phenomenon.

For a convincing argument that the case under inquiry is paradigmatic, paradigmatic cases should show that (1) the research program's core concepts, standards of observation, or levels of analysis are logically incoherent, and (2) the case is decisive for the sociological development of the research program. This approach combines some aspects of Lakatos's ([1978] 1980) methodology—in particular, the first-order core and the second-order belt heuristic—with Kuhn's ([1962] 1970) emphasis on the sociological origins of paradigmatic shifts. A straightforward example from the research program surrounding the emergence of popular democracy illustrates how these characteristics may interact. Although democracy researchers often associate the development and diffusion of politically competitive regimes with a small universe of cases in Western Europe and North America, research demonstrates that these new regimes devised multiple, overlapping patterns of hierarchy to exclude different members from the democratic polity (Caraway 2004). Logically, the development of these nascent democratic regimes co-existed with social hierarchy and inequality. Research about sociologically decisive cases that have steered this program—for example, the United Kingdom and the commonwealth colonies—shows that exclusion by race, gender, colonial subjecthood, and citizenship was central to early concepts of democracy among those who gained access to the franchise and free political participation (e.g., Holt 1992). Inquiry into these cases demonstrates that these patterns of social hierarchy and inequality merit systematic attention from researchers advancing the program about democracy's emergence.

These two standards of paradigmatic-case evidence are mutually necessary. Without demonstrating that the case renders the program logically incoherent, the researcher's empirical work provides an intellectual or political history of an influential case but not one on which the core assumptions of the research program hinge. Accounts of neoliberal political economy, for example, often refer to the influence of international development economists in the so-called "Chicago School." The Chicago School economists viewed the Augusto Pinochet government in Chile as a proving ground for aggressive pro-market liberalization policies. Although these international policy networks were prominent in the Chile case, Etchemendy (2011) demonstrates that Chile was one among several concurrent cases in which Chicago School economists sought out different forms of economic experimentation. For this reason,

disconfirming evidence of the neoliberal model of development policy in Chile does not have paradigmatic implications. Evidence that Chilean firms, domestic policymakers, and international development economists converged on the same policy outcomes falsifies the rigid neoliberal model *in the Chilean context*, but without broader implications for the research program (Silva 1996; Kurtz 1999).[7] The neoliberal model may persist in the Chicago School's broader policy advocacy in other countries, despite the Chile case's influence in the conceptual development and empirical study of neoliberalism.

The reverse—incoherence without evidence of the case's unique resonance among researchers—may chip away at Lakatos's "protective belt" of second order implications but fall short of undermining the research program's core. Early empirical research about genocide, for example, associated this type of large-scale violence with prerequisites of totalitarian rule and ideologies of wholesale destruction (e.g., Kuper 1981; Rummel 1995; Straus 2007). A proliferation of comparative studies during the late Cold War period has meant that multiple cases with divergent characteristics may qualify as the field-defining genocide, despite the common characterization of the Nazi Holocaust as an influential "index" case (Kopstein, Subotić, and Welch 2023; King 2012). Without an obvious case to anchor the development of a research consensus about the core properties of genocide, this core assumption remains an important—if contested—part of the universe of necessary conditions for how this violence emerges. This remains the case despite substantial evidence of resistance and rescue against these nominally-totalitarian conditions in important cases such as the Holocaust (Finkel 2017; Braun 2019; Einwohner 2003) and the Rwandan genocide (Fox and Nyseth Brehm 2018; Luft 2015), as well as evidence that genocide occurs in various different types of regimes. The Holocaust and the Rwandan genocide may be influential archetypes, but they fall short as "paradigmatic cases" because neither is a decisive empirical model for how and under what circumstances genocide may occur.

Cases may become paradigmatic through multiple different pathways. Because the sociological process of consensus-making leads researchers to emphasize some specific cases over others, different structures of collective inquiry may lead to different types of cases. I suggest a three-part continuum of the relationship between research and the political practice by which researchers reach consensus around their programs: (1) rigid academic hierarchy; (2) an iterative relationship between academic inquiry and political practice; and (3) a fully permeable boundary between inquiry and practice, in which academic researchers are themselves practitioners. A research program characterized by academic hierarchy might emphasize a case that dominates the relevant academic canon, such as the above-mentioned British democracy.[8] In these instances, studies of specific cases by dominant works or researchers provide a center of gravity that researchers seek to challenge. These cases are less identifiable in programs where areas of research emphasis are less centralized, and the canon of influential works is less rigid.

Cases in the academic canon are less dominant in the other two pathways. Programs that entail the occasional interaction between research and practice might focus on cases that influence contemporary policy consensus, such as the above-mentioned Munich case (Khong 1992). The third category of paradigmatic cases, by contrast, might emphasize cases that informed both the scholarship and the applied experiences of scholar-practitioners. A systematic comparison of these different disciplinary conditions is beyond the scope of this short essay; further research might evaluate the extent to which these different types of programs align with different configurations of paradigmatic cases.

It is theoretically possible for multiple paradigmatic cases to emerge from the multiple research communities that contribute to empirical inquiry about a single topic. Even where academic communities are hierarchical—the circumstance in which we might expect a single paradigmatic case to be most likely—the hierarchy of the academic canon is never so rigid as to crowd out multiple approaches that draw on multiple foundational examples. As the genocide research program suggests, there is also a fourth, null possibility: that the research program fully lacks a paradigmatic case or set of cases. In the instance of genocide research, the program lacks a paradigmatic case because multiple instances of genocide contribute to common sources of consensus about the area of inquiry.

## Conclusion

In this essay, I aimed to introduce provisional standards for evaluating paradigmatic cases and their implications for research programs in political science. Paradigmatic cases may catalyze changes in a research program if they address at least one of the three core components of a research program: (1) its necessary conditions; (2) its standards of observation; or (3) its levels of analysis. In my own research program about pogrom violence, for example, a paradigmatic case of mass mobilization without evidence of centralized

---

7 I thank Sebastián Etchemendy for his thoughtful reflections about this example.
8 I thank an anonymous referee for this suggestion.

organization might shift the central focus of inquiry about pogroms from the top-down organization of violence to the bottom-up activation of pogrom participants (Brass 1996). In addressing the second standard, this case could also emphasize that the spatial diffusion of violence is more central to understanding pogroms than the central decisions of pogrom organizations. Further still, this case might shift empirical attention to the individual or small-group factors that motivate violent participation.

Each of these potential changes may occur because research about the paradigmatic case demonstrates that these factors have received insufficient attention from the researchers that make up the program. In Lakatosian ([1978] 1980) terms, the paradigmatic case leads researchers towards a "progressive problemshift" that offers a more comprehensive explanation of the program's central phenomena than its predecessor. The existence of these logical standards for a paradigmatic case, however, does not itself guarantee that the research program's core assumptions will change. Like the case itself, that change is a by-product of the communities that produce and regulate scholarship around the program. These changes may occur from within the academic research community or because of the broader political or social environment that its research program inhabits.

This common vocabulary of paradigmatic-case research makes two main contributions to political science scholarship. First, these standards of evidence provide a necessary minimum of transparency that allows researchers within and outside of the discipline's many research programs to evaluate a case's potential implications for the paradigm under study. Without these standards, the multiple, sometimes conflicting standards by which researchers classify cases as "paradigmatic" will persist.

Second, these standards increase the likelihood that a clearer methodology of scientific research programs—the systematic evaluation of advancements in scientific discovery—may be useful in evaluating new knowledge about social and political behavior. This is an admittedly high bar. In future work, researchers should assess the potential pathways through which paradigmatic cases emerge and their intellectual contributions to specific scientific research programs. Absent this systematic inquiry, however, a common vocabulary around the empirical base on which these paradigms stand aids the progress of knowledge in political science by defining the boundaries of our discipline's research programs and identifying how they may change.

## References

Balcells, Laia. 2010. "Rivalry and Revenge: Violence against Civilians in Conventional Civil Wars." *International Studies Quarterly* 54, no. 2 (June): 291–313. https://doi.org/10.1111/j.1468-2478.2010.00588.x.

Beck, Robert J. 1989. "Munich's Lessons Reconsidered." *International Security* 14, no. 2 (Fall): 161–91. https://doi.org/10.2307/2538858.

Bellin, Eva. 2012. "Reconsidering the Robustness of Authoritarianism in the Middle East: Lessons from the Arab Spring." *Comparative Politics* 44, no. 2 (January): 127–49. https://doi.org/10.5129/001041512798838021.

Berry, Andrew. 2014. *Alfred Russel Wallace: A Short Biography*. Cambridge: Harvard University Press. https://doi.org/10.4159/harvard.9780674416468.c1.

Brass, Paul R., ed. 1996. *Riots and Pogroms*. New York: Springer.

Braun, Robert. 2019. *Protectors of Pluralism: Religious Minorities and the Rescue of Jews in the Low Countries during the Holocaust*. New York: Cambridge University Press.

Caraway, Teri L. 2004. "Inclusion and Democratization: Class, Gender, Race, and the Extension of Suffrage." *Comparative Politics* 36, no. 4: 443–60. https://doi.org/10.2307/4150170.Chandra, Kanchan, ed. 2012. *Constructivist Theories of Ethnic Politics*. New York: Oxford University Press.

Croke, Kevin, Guy Grossman, Horacio A. Larreguy, and John Marshall. 2016. "Deliberate Disengagement: How Education Can Decrease Political Participation in Electoral Authoritarian Regimes." *American Political Science Review* 110, no. 3 (August): 579–600. https://doi.org/10.1017/S0003055416000253.

Darwin, Charles. 1859. *On the Origin of Species: By Means of Natural Selection*. London: John Murray.

Eckstein, Harry. 1966. *Division and Cohesion in Democracy: A Study of Norway*. Princeton, NJ: Princeton University Press.

Eckstein, Harry. 1975. "Case Studies and Theory in Political Science." In *Handbook of Political Science*, edited by F. Greenstein and N. Polsby, 79–138. Boston: Addison-Wesley.

Einwohner, Rachel L. 2003. "Opportunity, Honor, and Action in the Warsaw Ghetto Uprising of 1943." *American Journal of Sociology* 109, no. 3 (November): 650–75. https://doi.org/10.1086/379528.

Etchemendy, Sebastián. 2011. *Models of Economic Liberalization: Business, Workers, and Compensation in Latin America, Spain, and Portugal*. New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511835223.

Finkel, Evgeny. 2017. *Ordinary Jews: Choice and Survival during the Holocaust*. Princeton: Princeton University Press.

Flyvbjerg, Bent. 2006. "Five Misunderstandings About Case-Study Research." *Qualitative Inquiry* 12, no. 2 (April): 219–45. https://doi.org/10.1177/1077800405284363.

Fox, Nicole, and Hollie Nyseth Brehm. 2018. "'I Decided to Save Them': Factors That Shaped Participation in Rescue Efforts during Genocide in Rwanda." *Social Forces* 96, no. 4 (June): 1625–48. https://doi.org/10.1093/sf/soy018.

Garand, James C., Micheal W. Giles, André Blais, and Iain McLean. 2009. "Political Science Journals in Comparative Perspective: Evaluating Scholarly Journals in the United States, Canada, and the United Kingdom." *PS: Political Science & Politics* 42, no. 4 (October): 695–717. https://doi.org/10.1017/S1049096509990205.

George, Alexander L., and Andrew Bennett. 2005. *Case Studies and Theory Development in the Social Sciences*. Cambridge, MA: MIT Press.Gerring, John. 2004. "What Is a Case Study and What Is It Good For?" *The American Political Science Review* 98, no. 2 (May): 341–54. https://doi.org/10.1017/S0003055404001182.

———. 2016. *Case Study Research: Principles and Practices*. New York: Cambridge University Press.

Giraudy, Agustina, Eduardo Moncada, and Richard Snyder. 2019. *Inside Countries: Subnational Research in Comparative Politics*. New York: Cambridge University Press.

Goertz, Gary. 2006. *Social Science Concepts: A User's Guide*. Princeton: Princeton University Press.

Grant, Peter R., and B. Rosemary Grant. 2002. "Unpredictable Evolution in a 30-Year Study of Darwin's Finches." *Science* 296, no. 5568 (April): 707–11. https://doi.org/10.1126/science.1070315.

Hathaway, Oona A. 2007. "Why Do Countries Commit to Human Rights Treaties?" *Journal of Conflict Resolution* 51, no. 4 (August): 588–621. https://doi.org/10.1177/0022002707303046.

Holt, Thomas C. 1992. *The Problem of Freedom: Race, Labor, and Politics in Jamaica and Britain, 1832-1938*. Baltimore, MD: JHU Press.Jackson, Patrick Thaddeus, and Daniel H. Nexon. 2009. "Paradigmatic Faults in International-Relations Theory." *International Studies Quarterly* 53, no. 4 (December): 907–30. https://doi.org/10.1111/j.1468-2478.2009.00562.x.

Khong, Yuen Foong. 1992. *Analogies at War: Korea, Munich, Dien Bien Phu, and the Vietnam Decisions of 1965*. Princeton: Princeton University Press.

King, Charles. 2012. "Can There Be a Political Science of the Holocaust?" *Perspectives on Politics* 10, no. 2 (June): 323–41. https://doi.org/10.1017/S1537592712000692.

Koch, Tom, and Kenneth Denike. 2009. "Crediting His Critics' Concerns: Remaking John Snow's Map of Broad Street Cholera, 1854." *Social Science & Medicine* 69, no. 8 (October): 1246–51. https://doi.org/10.1016/j.socscimed.2009.07.046.

Kopstein, Jeffrey S., Jelena Subotić, and Susan Welch. 2023. *Politics, Violence, Memory - The New Social Science of the Holocaust*. Ithaca: Cornell University Press.

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

Kuper, Leo. 1981. *Genocide: Its Political Use in the Twentieth Century*. New Haven: Yale University Press.

Kurtz, Marcus J. 1999. "Chile's Neo-Liberal Revolution: Incremental Decisions and Structural Transformation, 1973–89." *Journal of Latin American Studies* 31, no. 2 (May): 399–427. https://doi.org/10.1017/S0022216X99005325.

Laipson, Ellen. 2011. "Seismic Shift: Understanding Change in the Middle East." Stimson Center, Washington, DC.

Lakatos, Imre. (1978) 1980. *The Methodology of Scientific Research Programmes: Volume 1: Philosophical Papers*. New York: Cambridge University Press.

Luft, Aliza. 2015. "Toward a Dynamic Theory of Action at the Micro Level of Genocide: Killing, Desistance, and Saving in 1994 Rwanda." *Sociological Theory* 33, no. 2: 148–72. https://doi.org/10.1177/0735275115587721.Manekin, Devorah. 2013. "Violence Against Civilians in the Second Intifada: The Moderating Effect of Armed Group Structure on Opportunistic Violence." *Comparative Political Studies* 46, no. 10 (October): 1273–1300. https://doi.org/10.1177/0010414013489382.

Mills, Albert J., Gabrielle Durepos, and Elden Wiebe. 2010. *Encyclopedia of Case Study Research*. Thousand Oaks: SAGE.

Popper, Karl R. (1934) 2002. *The Logic of Scientific Discovery*. London: Psychology Press.

Rummel, R. J. 1995. "Democracy, Power, Genocide, and Mass Murder." *The Journal of Conflict Resolution* 39, no. 1 (March): 3–26. https://doi.org/10.1177/0022002795039001001.

Sa'adah, Anne. 2006. "Regime Change: Lessons from Germany on Justice, Institution Building, and Democracy." *Journal of*

*Conflict Resolution* 50, no. 3 (June): 303–23. https://doi.org/10.1177/0022002706286956.

Silva, Eduardo. 1996. "From Dictatorship to Democracy: The Business-State Nexus in Chile's Economic Transformation, 1975-1994." *Comparative Politics* 28, no. 3 (April): 299–320. https://doi.org/10.2307/422209.

Snow, John. 1855. *On the Mode of Communication of Cholera*. London: John Churchill.

Straus, Scott. 2007. "Second-Generation Comparative Research on Genocide." *World Politics* 59, no. 3 (April): 476–501. https://doi.org/10.1017/S004388710002089X.

Treisman, Daniel. 2020. "Economic Development and Democracy: Predispositions and Triggers." *Annual Review of Political Science* 23, no. 1 (May): 241–57. https://doi.org/10.1146/annurev-polisci-050718-043546.

Van Evera, Stephen. 1997. *Guide to Methods for Students of Political Science*. Ithaca, NY: Cornell University Press. Walker, Thomas C. 2010. "The Perils of Paradigm Mentalities: Revisiting Kuhn, Lakatos, and Popper." *Perspectives on Politics* 8, no. 2 (June): 433–51. https://doi.org/10.1017/S1537592710001180.

Worboys, Michael. 2000. *Spreading Germs: Disease Theories, and Medical Practice in Britain, 1865-1900*. New York: Cambridge University Press.

# Original Article

# Collaboration as Decolonization? Methodology as a Framework for Research with Indigenous Peoples

Mneesha Gellman[1]
*Emerson College*

## Introduction: Trust as Baseline

On a bright afternoon in March 2022, Victoria Carlson, the Yurok Language Program Manager for the Yurok Tribe stood in front of Yurok language students at Hoopa Valley High School, on the Hoopa Valley Indian Reservation in far Northern California. As she invited students to take home the informed consent permission forms I distributed through the maze of desks, she told students about how our collaboration started. "I didn't know what to think of Mneesha at first. You know, there have been a lot of researchers who come to Native communities and don't do right by us. But she kept coming back year after year and kept asking us what we thought and what we wanted.

And now, it is really exciting to see what our research together looks like." Victoria's comment, showing rightful skepticism of the outsider replaced over time by mutual appreciation for partnership, stays with me. It is part of the story of how I came to do what I term collaborative methodology, an explicit attempt to decolonize political science research.

I have spent the last two decades doing research in and with Indigenous communities, first in Latin America and more recently in far Northern California. Over time, it became clear to me that the "traditional" mode of entering a community to gather data, theory-test, and then scurry back to the office to write up findings in academic publications was not acceptable. In fact, given

the history of exploitation of Indigenous communities by researchers and everyone else (Deloria Jr. [1969] 1988, Norton 1979), such an extractivist model furthers a neocolonial dynamic. This was the opposite of what I aspired to do in my research.

Many sorts of methods, meaning the technical tool kit of data collection techniques, can be invoked in different methodologies, which are philosophies or frameworks that guide the purpose of how methods are conceptualized and implemented. While some methods may be more extractive than others, I focus here on the methodological—philosophical—orientation of the overarching research program rather than the specific tools of data collection. I use this focus because there is the possibility to make any method of data collection more or less extractive or collaborative and doing so hinges on the research methodology or framework; thus, my focus is on the overarching research program.

Doing fieldwork that is explicitly rooted in a collaborative framework inclusive of stakeholders has become a core part of my code of ethics as a researcher. This means that I work with communities that are looking for opportunities to answer questions that my social science research skill set can serve, while inviting people most affected by the research themes to lead the framing of research questions and define the methods used to answer them. To share what the collaborative methodology process can look like, this short essay offers a brief reflection from my own research and identifies obstacles that I have worked through as a collaborative methodologist.

## Paying Respect to Previous Scholarship

Indigenous scholars and allies are loudly and continuously calling for decolonization of research (Jacob 2013; Lee and Evans 2021; Mallon 2011; Tuhiwai Smith 2012; Wilson 2008). Such work needs to be done authentically and not metaphorically (Sheoran Appleton 2019; Tuck and Yang 2012). Particularly in education politics—which I engage by examining schools as spaces of citizenship production and youth identity formation—Indigenous scholars and others have longstanding critiques about the educational institutions that perpetuate neocolonialism and biases in research that supports White[2]-focused systems (Cleary and Peacock 1998; Jacob and RunningHawk Johnson 2020; Lara-Cooper 2017; Proudfit and Myers-Lim 2017). As a White, academically privileged outsider in the communities in which I work, I am also grateful to colleagues who have put into words the need for reflexive openness (Jacobs and Büthe et al. 2021, 188-9; Thomson 2021), which pushes us to center

positionality in the research process and to articulate how it impacts what we do and find. When I reflect on my own identity as a researcher in Indigenous communities, I find myself supported by structural privilege that must be addressed not as a sidenote but as a fundamental part of research design.

My contribution to these discussions has been to define collaborative methodology as a distinct enterprise and to show how researchers in the social sciences can maintain heralded values such as objectivity and transparency just as well in collaborative methodology as with any other framework (Firchow and Gellman 2021; Gellman 2021). To summarize these previous publications, collaborative methodology shares characteristics with participatory action research (PAR) and feminist methodology in that it directly addresses power dynamics and centers positionality. However, it differs from PAR in that collaborative methodology does not require action to flow from the research unless community stakeholders desire such action, and even then, researchers are there to support community efforts with citable research findings rather than direct action implementation. Collaborative methodology differs from feminist methodology in that it does not necessarily take a gendered perspective unless desired by community stakeholders, and it is different from member checking in that collaborative methodology is implemented throughout the life of a research project rather than only at the end when findings are shared back to ascertain accuracy.

Regarding accuracy and transparency, collaborative methodology has, in my experience, only made the work more accurate and transparent because I am being checked externally in my assumptions every step of the way when I am working with people affected by the research. From co-designing research questions and survey instruments to reading interview excerpt analyses out loud to interviewees, each step of the collaborative research process entails talking with people not solely as research subjects, but as partners in the investigatory process. Doing research this way can sharpen the focus of what is being researched, and it can enhance the quality of the research process itself.

## Identifying and Navigating Obstacles to Collaborative Methodology

Collaborative methodology has some drawbacks, and isn't necessarily something that can work for everyone. In the course of presenting collaborative methodology at conferences and workshops, I have had insightful conversations with colleagues who point out numerous

---

2 I follow the thinking that capitalizing White reduces its treatment as a default backdrop to normalcy and instead flags it as a particular identity, with structural implications, in line with the thinking outlined by Mack and Palfry (2020).

concerns.[3] Resource constraints—including time and money— along with career benchmark goals, are legitimate reasons to carefully evaluate how collaborative methodology is applicable in one's own scholarly life. I've written elsewhere about how collaborative methodology may not be a good fit for graduate students on compressed timelines, or for those who work with elites such as elected officials, where researchers may actually be less empowered than their participants (Gellman 2023; 2023 under contract). I've also noted that moving the research ahead can happen more slowly when one needs to wait for emails or phone calls from busy community partners who have other day jobs.

Yet, I've made it work. As a faculty member at a teaching college with a significant course load and modest research funding, I have persistently written grants to enable me to drag my family to research sites for semesters at a time. With careful advance planning, I have made collaborative methodology the center of my research, even when both time and money are scarce. It may not be a glamourous way to do research, but it can be a possible and effective way even with a shortage of funding and time.

For those of us who research topics relevant to historically and contemporarily marginalized communities, considering the role of collaboration acts as a baseline rather than a distant goal. Evaluating research designs to address the neocolonial power dynamics of extraction is part of ethical research. Resource constraints for scholars are simply not a sufficient reason to perpetuate an extractive intellectual industry. It is our responsibility to find ways around and through the constraints.

While I urge us to overcome obstacles and engage collaboratively, I recognize that many marginalized communities in which scholars may want to work do not have clear or consensual leadership structures through which to define collaborative process. The reality of leadership fragmentation poses a particular challenge and can be addressed in multiple ways. It may be necessary to work with multiple organizations, for example, layering questions and methods from more than one community partner into a research puzzle. Alternatively, researchers can clearly define who they identify as community stakeholders and justify a more limited scope condition for community leadership, or some fusion of both practices may be most fruitful. In working with the Yurok Tribe, I entered a public community space with clearly defined leadership hierarchies, where the collaborative parameters of relationships are derived from elections, hiring practices, and seniority. Though there is fragmentation in many social groups, defining with whom one is working and why, as well as who is left out, can be the basis for successful collaborative research. I turn to specifics of the collaborative process in the section below.

## Research Framework: A Look Under the Hood

In 2016, I first sought and was granted permission from the Yurok Tribal Council to initiate a study on the impact of Yurok language access in two public high schools in far Northern California. I had been in conversation with the Yurok Tribe's Education Department in the year prior, and we had several discussions to outline the research so that it could be mutually useful. The Tribe wanted concrete information about the impact of Yurok language classes in the high schools, which they were partially funding, and I was interested in the connection between language access, youth identity formation, and participation choices. We talked through a research puzzle that could include both of our angles of interest, each making compromises to make the partnership work. For example, the Yurok Tribe at one point would have preferred a more in-depth analysis of students' academic records to understand grades and attendance in relation to Yurok language-enrolled students. I wasn't able to take that on because it was challenging enough to get guardian permission to do less invasive study components like surveys and focus groups. But I did agree to include a quantitative survey component in my otherwise qualitative toolkit, both to expand the number of students from whom I could gather information as well as to ensure a standardized set of closed-answer questions.

I also agreed to the stipulation that all research could only be published after first being reviewed by the Office of the Tribal Attorney and then being subject to approval for publication by the Yurok Tribal Council. This was of particular concern for the Tribe because of the long history of misrepresentation of Native people by outside researchers. In addition, it was not clear at the outset of this project if the findings would show a positive benefit to students having access to the Yurok language, and there was rightful concern by the Council that negative findings could harm language revitalization efforts. I had sufficient anecdotal evidence from preliminary research both on the ground and in the literature that the hypotheses—Indigenous language access does positive things for youth identity in different ways depending on youth demographics—that I was willing to take this risk. While I was not yet tenured when I began the project,

---

3 I particularly thank participants in the 2019 and 2021 Southwest Multi Method Research Workshops for rich conversations on the benefits and limitations of collaborative methodology.

this was my second book project. My first book was already in print, so my own career benchmarks were not as perilous as with a first project. Nevertheless, familiarity with an issue and trust within a community are clearly assets that make agreeing to pre-publication review and approval possibly more successful.

Agreeing to have all publications reviewed by the Tribe meant that I as the researcher gave up a significant amount of autonomy to bring this research project to fruition. Radically breaking from political science tradition, the research is not solely mine. I am only one of its originators and stewards. By agreeing to the collaborative terms outlined by the Tribal Council, the research belongs to the Yurok Tribe as well, and we must work together to produce meaningful work that addresses both of our agendas.

For the sake of bare honesty, I will say that taking such risks as a researcher is terrifying. There was every possibility that I could have invested years of work into something that was ultimately not deemed publishable by the Tribe. My investment of time, grant money, and hopes for my own career goals might not have paid off. Any collaborative researcher may find themselves in this position. Yet, what actually happened has made the research more meaningful to stakeholders themselves. The Office of Tribal Attorney read my entire book manuscript as well as several related articles, including this one, and offered numerous insights, through queries or comments, that made the works better and more accurate before they became public facing. Had I attempted to do the research without the collaborative relationship with the Tribe, not only might my finished work have contained avoidable errors, but I also would have risked angering and alienating the Yurok Tribe by talking about something in their purview without asking their permission and working with them. At its most basic, engaging stakeholders in research design and implementation through collaborative methodology is a way to show respect for the autonomy of sovereign communities, something that is fundamental to working with Indigenous communities. In addition, collaborative methodology facilitates local ownership over research, which can be another element in decolonizing research.

To those who are concerned that the intimacy of collaboration with one's research community introduces the potential for bias, remember that bias is no more plausible in collaboration than in extractive research. In our co-edited symposium on collaborative methodology, peace scholar Pamina Firchow and I make this point, together with other contributors in their own articles (Firchow and Gellman 2021). Like all good scholarship, careful scope conditions, reaching saturation in data collection, triangulation of findings, and constant exploration of intervening variables can all be present in collaborative methodology.

It is worth mentioning that a mixed methods tool kit may be especially useful in collaborative work because it expands the number of ways that indicators can be documented as data and checked against each other. This can be an important way to share with community partners the depth of the findings. Such multi-methods work also lends itself to being used and applied by community partners in a broader range of ways. In my own work, I used both qualitative and quantitative data collection methods, including ethnography, interviews, focus groups, and surveys in a multi-sited comparative project that spanned into four public high schools across Oaxaca, Mexico and far Northern California. The combination of mixed methods and comparative design set me up to have a high level of data points in what was ultimately a community-based small n study. I recently published a large portion of the data in my newest book, *Indigenous Language Politics in the Schoolroom: Cultural Survival in Mexico and the United States* (2023).

The collaborative methodology approach can contribute to sustaining relationships between researchers and stakeholders. In 2021, while writing the manuscript, the Yurok Tribe Education Department invited me to help conceptualize what we informally call Phase Two, the next iteration of our research together. While the *Indigenous Language Politics* book is comparative across both California and Oaxaca, it only covers two of the four high schools where Yurok is taught in California. The Yurok Tribe's Yurok Language Program was interested in the connection between Yurok language classes and resilience, with a focus on both Covid and more broadly, the high level of intergenerational trauma that Native youth navigate. I continue to be interested in the role that educational curriculum plays for minority students from a range of backgrounds. In a series of conversations via Zoom as well as dozens of emails and draft exchanges, we developed new questions on these themes and on the effect that culturally connected curricula might have on students of the Yurok language. Into this research I added in my additional interests in Latinx identity by including classes such as Spanish for Spanish-speakers and English Language Development in a comparative study design.

In December 2021, the Yurok Language Program and I brought our research proposal to the Yurok Tribal Council, where it was approved (almost but not quite unanimously). I next completed my institution's IRB, and I secured leave from Emerson College via a grant. In Spring 2022, I moved my family to far Northern California, and set up numerous meetings to secure permission from four different school district

superintendents and four school principals, and I also met with and obtained permission from the teachers themselves to run the study in their classes. I spent four months driving from one high school to the next, texting and emailing teachers to confirm the next day's visit, and begging office space from front desk staff to conduct interviews and focus groups with a modicum of privacy.

Across all the research instruments, we had co-designed questions on topics such as ways of coping during the pandemic year of online school, sources of support for youth wellbeing, and misrepresentation of Indigenous and other minority peoples in the K-12 curriculum. This focus on resilience—what wears it down and what builds it up—is helping bring to light the complex identity issues that high school students encounter when faced with a culturecidal curricula that prioritize White narratives. In these examples, taking the time to build trust and work with community stakeholders for a mutually beneficial research design led to a successful research experience for not one but two book-length projects. At the end of the day, the Yurok Tribe will have sound data documenting the impact of some of the cultural work in which it has invested, and I will have publications that further my academic career while also holding the potential to directly inform education policy at the community level.

## Conclusion

Defining the parameters of our research methodologies is a deeply political act. Deciding who is a "subject" and who is a stakeholder or principal investigator reveals a power hierarchy that can further deepen the dynamics of settler colonialism. The social sciences, including political science, can serve a noble role in society by helping us to better understand the world around us. If political scientists keep using tools of domination to reach for that understanding, we risk embedding frameworks of colonization into the research we produce.

Collaborative methodology is one approach that can shift the paradigm. It is not an easy philosophy to embrace for some researchers because it requires overcoming numerous resource obstacles, and it requires taking the risk of vulnerability in sharing ownership over research process and products. Yet, the benefits to collaboration are profound. The relationships I have formed with colleagues in the Yurok Education Department rest on a bedrock of mutual appreciation and trust that deepens the worth of the findings themselves.

I did not have Victoria's implicit trust when we first met because I come from a scholarly discipline and positionality that has caused her community harm in myriad ways since colonization. By committing time and resource to working together, we found ways to ask questions and obtain answers that meet multiple goals. As we analyze the world's problems and try to find ways to reduce harm and address structural violence, collaborative methodology is a way to start within our discipline.

## References

Cleary, Linda Miller, and Thomas D. Peacock. 1998. *Collected Wisdom: American Indian Education*. Boston: Allyn and Bacon.

Deloria Jr., Vine. (1969) 1988. *Custer Died For Your Sins: An Indian Manifesto*. Norman: University of Oklahoma Press.

Firchow, Pamina, and Mneesha Gellman. 2021. "Collaborative Methodologies: Why, How, and for Whom?" *PS: Political Science & Politics* 54, no. 3 (July): 525-9. https://doi.org/10.1017/S1049096521000330

Gellman, Mneesha. 2021. "Collaborative Methodology with Indigenous Communities: A Framework for Addressing Power Inequalities." *PS: Political Science and Politics* 54, no. 3 (July): 535-8. https://doi.org/10.1017/S1049096521000299.

Gellman, Mneesha. 2023. *Indigenous Language Politics in the Schoolroom: Cultural Survival in Mexico and the United States*. Philadelphia: University of Pennsylvania Press.

Gellman, Mneesha. (2023, under contract). "Considering Collaboration as Part of Your Research Design." In *Doing Good Qualitative Research,* edited by Jennifer Cyr and Sara Wallace Goodman. Oxford: Oxford University Press.

Jacob, Michelle M. 2013. *Yakama Rising: Indigenous Cultural Revitalization, Activism, and Healing*. Tuscon: University of Arizona Press.

Jacob, Michelle M., and Stephany RunningHawk Johnson, eds. 2020. *On Indian Ground: A Return to Indigenous Knowledge-Generating Hope, Leadership and Sovereignty through Education*. Charlotte: Information Age Publishing.

Jacobs, Alan M., and Tim Büthe et al. 2021. "The Qualitative Transparency Deliberations: Insights and Implications." *Perspectives on Politics* 19, no. 1 (March): 171-208. http://doi.org/10.1017/S1537592720001164.

Lara-Cooper, Kishan. 2017. "Protecting the Treasure: A History of Indigenous Education in California." In *On Indian Ground: A Return to Indigenous Knowledge—Generating Hope, Leadership and Sovereignty through Education*, edited by Joely Proudfit and Nicole Quinderro Myers-Lim, 1-18. Charlotte: Information Age Publishing.

Lee, Emma, and Jennifer Evans, eds. 2021. *Indigenous Women's Voices: 20 Years on from Linda Tuhiwai Smith's Decolonizing Methodologies*. London: Bloomsbury Academic.

Mack, Kristen, and John Palfrey. 2020. "Capitalizing Black and White: Grammatical Justice and Equity." *MacArthur Foundation*, August 26, 2020. https://www.macfound.org/press/perspectives/capitalizing-black-and-white-grammatical-justice-and-equity.

MacLean, Lauren M., Elliot Posner, Susan Thomson and Elisabeth Jean Wood (2021). "Research Ethics and Human Subjects: A Reflexive Openness Approach." In The Qualitative Transparency Deliberations: Insights and Implications (QTD Steering Committee report, eds. Alan Jacobs and Tim Büthe. *Perspectives on Politics*. 19, no.1 (March): 188-189.

Mallon, Florencia E., ed. 2011. *Decolonizing Native Histories: Collaboration, Knowledge, and Language in the Americas*. Durham: Duke University Press.

Norton, Jack. 1979. *Genocide of Northwestern California: When Our Worlds Cried*. San Francisco: Indian Historian Press.

Proudfit, Joely, and Nicole Quinderro Myers-Lim, eds. 2017. *On Indian Ground: A Return to Indigenous Knowledge—Generating Hope, Leadership and Sovereignty through Education*. Charlotte: Information Age Publishing.

Sheoran Appleton, Nayantara. 2019. "Do Not 'Decolonize' . . . If You Are Not Decolonizing: Progressive Language and Planning Beyond a Hollow Academic Rebranding." *Critical Ethnic Studies* (blog) February 4, 2019. http://www.criticalethnicstudiesjournal.org/blog/2019/1/21/do-not-decolonize-if-you-are-not-decolonizing-alternate-language-to-navigate-desires-for-progressive-academia-6y5sg

Thomson, Susan. 2021. "Reflexive Openness as Ethical Research Practice." *PS: Political Science & Politics* 54, no. 3 (July): 530-4. https://doi.org/10.1017/S1049096521000159.

Tuck, Eve, and Wayne K. Yang. 2012. "Decolonization is Not a Metaphor." *Decolonization: Indigeneity, Education & Society* 1 (1): 1-40. https://jps.library.utoronto.ca/index.php/des/article/view/18630.

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

Wilson, Shawn. 2008. *Research is Ceremony: Indigenous Research Methods*. Winnipeg: Fernwood Publishing.

# Original Article

# Overcoming the Laws-in-Translation Problem: Comparing Techniques for Translating Legal Texts

Anthony J. DeMattee
*Emory University*

Takumi Shibaike
*University of Calgary*

Nick Gertler
*Concordia University Montreal*

Elizabeth A. Bloodgood
*Concordia University, Montreal*

Statutes and regulatory rules (henceforth "legal text") are invaluable forms of primary data for comparative analysis of regime transitions, political participation, policy diffusion, and the realization of justice and rights. Governments carefully curate and archive their legal codes, making it possible to trace a law's history and track its diffusion to other jurisdictions. Governments do not consistently enforce laws as they are written (Pound 1910; Law and Versteeg 2013); to understand the distinction between laws in the books and laws in action, we first must be able to read the law. Legal texts are thus invaluable for historical analysis and comparative studies, as laws in one jurisdiction can have important similarities and differences across national contexts (Glasius, Schalk, and De Lange 2020; Hummel, Gerring, and Burt 2021; Berinzon and Briggs 2019).

Despite their importance and accessibility, legal texts are not studied in a comparative perspective as frequently as quantitative indicators because of the limitations of language. The European Union and United Nations are

leading efforts to collect and translate legal texts, but those collections are not comprehensive across topics, countries, or time (Giampieri 2016). We refer to this as the "laws-in-translation problem," a term that we use to refer to the fact that legal translations are scarce or incomplete, and available strategies to translate such texts are either cost-prohibitive or error-prone to an unknown degree.

Researchers can deploy solutions to the laws-in-translation problem, but each presents new issues. Professional human translators remain the "gold standard" for accuracy (Lucas et al. 2015, 259-60), but human translation services quickly become expensive for large legal codes. Alternatively, machine translation (MT) software provides speed, accessibility, and affordability, but there are concerns about their accuracy and reliability. We evaluate DeepL, Google, and Microsoft MT applications for their effectiveness in translating legal texts on five dimensions: generalizability, flexibility, presentation, simplicity, and reliability. For a controlled comparison of machine and human translations, we translate legal content from similar laws enacted in Brazil, China, France, Japan, and Mexico. We find that MT tools are not sufficiently nuanced for legal practice or fine-grained analysis, but maintain that MT's accuracy—when used in a hybrid approach—is sufficient for researchers conducting comparative socio-legal and policy research. Our recommendation to those who seek accuracy and cost-efficiency is to use MT applications in tandem with human translators. As we show in the following sections, combining these translation strategies increases transparency and accuracy while lowering costs and decreasing time spent.

Our hybrid approach to translating legal texts razes methodological barriers and expands the number of cases available for comparative analysis. The method has implications beyond law and policy scholars. The language in which a government writes its laws and executive orders has no relationship with whether it follows, bends, or violates those legal rules. Whether legal texts appear in Russian or English matters little if the government enforcing those rules seeks to threaten fundamental freedoms, undermine elections, or loot personal property. Yet language barriers profoundly hinder our ability to evaluate whether laws authorize illiberal practices, disguise them as legitimate actions, or simply fail to prevent such phenomena. Overcoming these barriers allows us to read local reports covering protests in foreign capitals and access the ideas and reactions of those affected by war, migration, or disaster.

## Motivation: The Missing Standard for Translating Legal Texts

The social sciences contain numerous examples of scholars who have successfully navigated the laws-in-translation problem. Unfortunately, published research often lacks a clear explanation of *how* the translation process unfolded. We are guilty of this ourselves (DeMattee 2022a; Bloodgood, Tremblay-Boire, and Prakash 2014). For research covering a multilingual legal corpus, readers must assume that the researchers are polyglots or that translation is a rigorous part of data collection and research design phases. Not disclosing the translation process is not the same as being unwilling to be transparent. The publishing process has limited authors' ability to communicate how they translated texts and verified translations. Innovative technology such as the Qualitative Data Repository's Annotation for Transparent Inquiry (ATI) eases some limitations. ATI enables researchers to digitally link their article to analytic notes, allowing authors to provide more information about the data and analytic choices, including extended excerpts of original and translated legal text (Kapiszewski and Karcher 2021; Elman, Kapiszewski, and Lupia 2018). Clear explanations and standards for the translation of laws increase the replicability and credibility of findings. Reviewing sociology, political science, and law and society journals reveals a surprising lack of comparative law and policy research across geographic regions and languages. This research gap is likely related to a missing process for effective and efficient translation. Solving the laws-in-translation problem is thus likely to open opportunities for research into new questions as well as new approaches to old puzzles.

One approach is to limit case selection to regions whose countries publish laws in a common language. Scholars have successfully compared the development of penal codes in French West African countries (Berinzon and Briggs 2019) and corporate law in China, Hong Kong, and Taiwan (Lin and Chang 2018). Similarly, Lemon and Antonov (2020) compare legal text in five post-Soviet countries to show that independent countries enact laws with nearly identical language. Researchers can encounter the laws-in-translation problem even when maintaining a regional focus. For example, when studying immigration laws in the Americas, Cook-Martín and FitzGerald (2019) needed to analyze a multilingual corpus that included countries colonized by Britain, France, Spain, and Portugal. While researchers might need to limit their scope to monolingual corpora or regional analyses if a deep and nuanced interpretation of the particular legal text is necessary, scholars who focus on global phenomena, including diffusion effects

and postcolonial change, need competencies in multiple languages or an alternative approach.

Secondary sources offer scholars one solution to expand their geographic and linguistic range. They provide a valuable foundation, and merging multiple sources can improve comprehensiveness. Glasius, Schalk, and De Lange (2020, 457) and Hummel, Gerring, and Burt (2021, 873) assemble their corpora from various public, private, and academic sources. Secondary sources have hidden costs as they are neither as comprehensive nor as accurate as researchers often require, especially if secondary sources are limited to specific languages or periods. Echoing previous caveats of off-the-shelf data (Bennett 2007), to what degree can researchers trust secondary sources' accuracy, rigor, and objectivity? The remedy suggested still applies: "Take between five and ten random observations from the dataset and attempt to code the variables *from the ground up*" (Goemans 2007, 12; emphasis added). This prescription to recode a random sample of legal texts returns us to the laws-in-translation problem.

Brute force is another pathway to obtaining a large, multilingual corpus. With adequate resources, researchers can collect, translate, and code primary sources from the ground up. With financial commitments from Google, the National Science Foundation, and the United States Institute of Peace, the Comparative Constitutions Project (Elkins and Ginsburg 2021; Elkins, Ginsburg, and Melton 2009) is the exemplar for overcoming the laws-in-translation problem. Bradford et al. (2019, 416) used the collective skills of 70 law school students over six years to code competition laws in 131 jurisdictions between 1889 and 2010. DeMattee (2020) spent 567 person-hours coding a six-language corpus of 285 laws enacted by seventeen countries between 1872 and 2019. Researchers who wish to embark on similar quests should not underestimate the financial and human resources necessary to execute such projects.

The relative lack of comparative legal data, combined with the costs in time and money to create it, demonstrate why it is vital to discover new processes to allow researchers to accurately and economically translate legal texts. Appropriately using MT applications opens legal research across languages for more innovative comparative studies and increases access to less common case studies. Defining best practices for using these tools is a new and valuable contribution to qualitative and mixed-methods research.

## Methods

To test the comparative performance of the MT applications against one another and human translators, we use similar legal texts from five countries that are written in major international languages. We examined three languages that use the Roman alphabet and two that use logograms. We used the legal definition of civil society organizations as specified in laws enacted by governments in Brazil, China, France, Japan, and Mexico to find comparable legal text across countries. These legal definitions vary slightly in content and length. Supplemental information (DeMattee et al. 2022) contains all original legal texts and the translated versions produced by DeepL, Google, and Microsoft.

We compared the original and translated versions of these legal definitions to assess which MT application has the highest usability and reliability across languages. First, we evaluated each translation application according to its flexibility in accepting and outputting file types, its ability to preserve the document's structure, and the available number of languages. Second, we tested the reliability and quality of the translations by testing the MT versions against human translations. Two native speakers per language of interest first translated the source text into English. Then, these same multilingual speakers evaluated four translated versions of a single source text: three MTs and one human. We randomized the order and anonymized the source of the translated texts and instructed evaluators to conduct their single-blind evaluations independently and without the assistance of other tools. Each evaluated the translated texts according to whether they required minor or critical edits for grammar and meaning. We defined minor edits as corrections made to the translated text that maintained the meaning of the source text, even if the translation is mediocre, and critical edits as corrections made to the text that did not maintain the original meaning.

### Defining Five Measures: Generalizability, Flexibility, Presentation, Simplicity and Reliability

We define five measures to evaluate the quality of MT applications based on our past experiences with comparative research on association and charity laws (DeMattee 2022a, b; Bloodgood, Tremblay-Boire, and Prakash 2014) These measures—generalizability, flexibility, presentation, simplicity, and reliability—vary in importance depending on the research project.

*Generalizability* refers to the number of languages available for text-to-text translation. MT applications may be available for transliteration, translating the text in images, or text-to-speech translation. We expect that generalizability will increase with time as the demand grows for automated translation within and across applications, particularly among smartphone users. Increased generalizability does not guarantee that other measures of translation effectiveness will improve at the same pace. We argue that it is equally important, if not

more so, to know which MT application provides the most accurate translation for a given language.

*Flexibility* reflects the compatibilities of file formats in MT applications. Limitations on flexibility come in three forms: intake format, file size, and output format. When files are not compatible with MT applications, researchers must first convert them into appropriate formats. Another consideration for flexibility is file size, as some MT applications limit a file's upload size. Finally, output format typically corresponds to input format, but some MT applications allow more options for output formats. Flexibility in output format allows researchers to choose appropriate formats for subsequent analysis, reducing the incidence of human error.

*Presentation* is the degree to which the MT application preserves the source text's formatting and layout. Preserving identifiers accurately (e.g., alpha or numeric ordering) allows researchers to reference sections of a law correctly while facilitating replicability. If the translation process suppresses or distorts identifiers, researchers may need to invest considerable energy re-identifying articles, sections, subsections, and paragraphs for correct citations. This risks introducing human errors. Maintaining formatting (e.g., alignment and hanging indentations) is another consideration. Consistent formatting allows researchers to easily navigate and compare the original and translated versions. Likewise, protecting page breaks is stylistically desirable and valuable when a translation application strips identifiers from the document.

*Simplicity* is the number of actions necessary to translate a single document, including preparing files to be readable in MT applications. This factor becomes more relevant as the number of files, or the size of the files, in the legal corpus increases. Greater simplicity means less work to prepare files prior to translation and fewer opportunities for mistakes. Some PDF files are document images or scans of computer-generated text. These situations require an additional step. Optical character recognition (OCR) software converts images into machine-encoded text. The original composition of the text (e.g., manual versus computer typesetting) and image quality can affect OCR detection. In both cases, researchers may benefit from specialized software to improve OCR detection and obtain better results.

*Reliability* represents accuracy. The most important aspect of reliability is that translation maintains the meaning of original texts. If the translated text's meaning differs from the original, any analysis that follows will be severely flawed. Another consideration for reliability is grammatical and syntax accuracy; however, such errors may be minor enough not to alter the meaning of the original texts.

We use two tests to evaluate *reliability*. Our first measure uses native language speakers to evaluate translations of civil society laws in five languages. For each language, two single-blinded evaluators independently assess the source text and translated text at three levels. "Minor edits" are small, stylistic changes that polish the text to improve flow or readability. "Critical edits" involve substantive changes necessary to realign the translated text with the source text. Substantive changes go beyond slight improvements in readability and correct key errors that may otherwise jeopardize research findings. Not all critical edits require large-scale changes. For example, incorrectly translating a deontic such as "may" for "must" can seriously impact research findings. Finally, each evaluator made an overall assessment of whether the translated text maintained the source text's original meaning without any editing.

The second reliability measure is a series of round-robin translations. Here, we translate the source text to other languages and then translate it back to the original language. Translation applications, we assume, treat each translation as an independent task. Errors will therefore compile through multiple translations. Reliability is the similarity between the source text and the final translation reverted to the original language. We use English and German—two languages outside those studied—to conduct the backwards translations. We vary the number of translations from one to two foreign languages to further test each translation application's stability. We use a similarity score to assess these backwards translations. Similarity scores are the percentages of words in the backwards translation that identically, nearly, or relatedly match the words in the source text.

### Overview of Machine Translation (MT) Tools

There are multiple applications available to researchers seeking automated translations. While their user interfaces may be similar, the algorithms that generate the translations vary. A number of human decisions feed into the final algorithms, which makes them objects of human creation capable of producing biased and fallible outputs (Diakopoulos 2013, 10; Salminen et al. 2020). Time and resources constrain programmers' ability to validate and update an algorithm's performance. This means that time and resources are additional factors that we expect will affect an MT application's accuracy. Older tools, or those maintained by organizations with greater resources, may be more accurate because of the greater availability of time and resources to train and debug the algorithms. Our research compares both free and proprietary algorithms provided by big and small companies, both old and new.

## Analysis

By our assessment, Google is the most *generalizable* MT application because it has the most languages available for text-to-text translation at 108. With 72 languages, Microsoft is the next most generalizable. Between these two tools, even researchers working in regional languages (e.g., Haitian Creole) have options for translation applications. All three applications offer eleven common languages: Chinese, Dutch, English, French, German, Italian, Japanese, Polish, Portuguese, Russian, and Spanish.

DeepL and Google Translate are the most *flexible* MT applications, and are the only translators that accept PDF files, which is the format used most often in legal texts. Microsoft Translator requires researchers to convert PDFs to another file type for translation. This rigidity in file formatting is not a critical factor when a legal corpus is composed of a small number of text documents, but could quickly become a significant issue. As the size of the legal corpus expands, researchers can use programs like R or Python to manage file conversions, but this adds new technical requirements. Applications also vary in their ability to process files of different size. DeepL can process bigger files and larger quantities of files if users upgrade to a premium service. Google Translate limits file size, but does not offer a subscription service to process large documents. Microsoft does not provide explicit file size limits on its Translator; however, our experience is that the Microsoft application tends to fail in files above 100 pages. Output file types are as inflexible as input types. Google Translate currently does not offer an export function; all translations are displayed as a webpage that researchers must save as a PDF. DeepL and Microsoft render translations as a new file that researchers can save in various file types.

DeepL and Microsoft slightly outperform Google on the *presentation* measure. While all three applications protect font styles, Google Translate is prone to dropping identifiers, such as alpha or numeric section markers that allow researchers to navigate a law's contents. Such omissions may force researchers to re-identify articles, sections, subsections, and paragraphs to accurately reference passages. Our experience working with large and multilingual legal texts raises two concerns. First, sections of laws may refer to one another or point to schedules for further information. This means that researchers may need to translate entire laws rather than individual sections, necessitating accurate section markers. Second, laws can be long and the ability to accurately translate a large document is a critical matter.

*Simplicity* varies depending on the quality of the original document. If the starting point is a corpus of DOCX files, the three tools (DeepL, Google Translate, Microsoft) have equal simplicity, as they do not require users to reformat files before translation. If the corpus contains machine-encoded PDF files, then DeepL, Google, and Microsoft require those files be converted to a compatible format, which varies across applications. If PDF files contain scanned images of text, researchers must first convert them to text before translating the document. Researchers can use external OCR programs, such as Adobe Acrobat and R, to automate the conversion processes. In our comparison, Google Chrome always requires one more step than the other applications because researchers must convert files to HTML format.

*Reliability* estimates the precision of each translation application and the degree to which it is a reliable research tool. We discuss two measures of reliability: human evaluators' assessments of the translations, and round-robin similarity scores. We begin with human evaluators who edited and analyzed the machine translations at three levels: minor edits, critical edits, and whether the translation maintained the meaning of the source text. The applications' average scores are similar concerning minor edits; as Table 1 shows, we found approximately four minor edits for every 100 words. Note that human translators amended other human translations at nearly twice the rate—seven minor edits per 100 words—as the machine translations. Our tests on civil society laws suggest that, for most research purposes, minor errors do not prevent researchers from understanding and using the translation. In many cases, minor errors simply require researchers to work through legal text that is wordy or convoluted.

Our remaining two measures are more consequential. Critical edits are substantive changes that are necessary to realign the translation with the source text. Critical changes correct errors that may jeopardize research findings. Google Translate, which averaged 0.5 critical edits per 100 words, was the top performer on this measure. DeepL and Microsoft translators produced over twice as many critical edits, on average. The third measure is whether a translation maintains a text's original meaning. Overall, native language speakers had perfect inter-coder reliability on this measure. Google Translate was the top performer in maintaining original meaning. Eight out of ten evaluators graded Google Translate positively on this measure (with the exception of the two Japanese assessments). Indeed, the Japanese language speakers agreed that all three translation applications drifted from the text's original meaning. The two Portuguese-speaking evaluators found that DeepL Translator likewise lost the text's original meaning, and the two Chinese-speaking evaluators observed a similar error in Microsoft Translator.

Table 1: Native Language Speaker Reliability Analysis

| | DeepL Translator | Google Translate | Microsoft Translator | Human Translators |
|---|---|---|---|---|
| **Brazilian Legal Text** | | | | |
| Minor Edits | 2.5 edits | 2.5 | 1.5 | 4.2 |
| Critical Edits | 2.5 edits | 0.8 | 0.8 | 0 |
| Maintained Meaning | No | Yes | Yes | Yes |
| Minutes to Edit Translation | 9.0 minutes | 4.1 | 4.5 | 5.1 |
| Words Added/Removed | +1.6 words | +2.5 | -2.3 | +2.6 |
| **Chinese Legal Text** | | | | |
| Minor Edits | 5.9 | 4.7 | 5.1 | 8.5 |
| Critical Edits | 0 | 0 | 1.7 | 0 |
| Maintained Meaning | Yes | Yes | No | Yes |
| Minutes to Edit Translation | 7.8 | 7.5 | 11.9 | 8.5 |
| Words Added/Removed | -2.9 | 0 | -11.9 | +12.3 |
| **French Legal Text** | | | | |
| Minor Edits | 5.6 | 3.6 | 1.2 | 11.5 |
| Critical Edits | 0 | 0 | 0 | 0 |
| Maintained Meaning | Yes | Yes | Yes | Yes |
| Minutes to Edit Translation | 3.3 | 3.6 | 4.8 | 8.7 |
| Words Added/Removed | +1.1 | 0 | +1.2 | +13.4 |
| **Japanese Legal Text** | | | | |
| Minor Edits | 0.4 | 0.8 | 2.3 | 3.4 |
| Critical Edits | 3.1 | 1.6 | 3.5 | 0 |
| Maintained Meaning | No | No | No | Yes |
| Minutes to Edit Translation | 6.6 | 12.3 | 8.74 | 8.5 |
| Words Added/Removed | -45.2 | +2.5 | -8.7 | 0 |
| **Mexican Legal Text** | | | | |
| Minor Edits | 6.4 | 8.1 | 9.0 | 6.4 |
| Critical Edits | 1.3 | 0 | 0 | 0 |
| Maintained Meaning | Yes | Yes | Yes | Yes |
| Minutes to Edit Translation | 19.2 | 41.9 | 21.8 | 18.9 |
| Words Added/Removed | +1.3 | +5.4 | +2.6 | +1.3 |
| **Five-Law Average** | | | | |
| Minor Edits | 4.2 | 3.9 | 3.8 | 6.8 |
| Critical Edits | 1.4 | 0.5 | 1.2 | 0 |
| Maintained Meaning | Yes (6/10) | Yes (8/10) | Yes (6/10) | Yes (10/10) |
| Minutes to Edit Translation | 9.2 | 13.9 | 10.3 | 9.9 |
| Words Added/Removed | -8.8 | +2.1 | -3.8 | +5.9 |

All measures were averaged across native language speakers (2 per country, 10 total). For comparability and interpretability, the measures of edits, minutes to edit, and words added or removed are all standardized to a common unit: per 100 words. Google Chrome was excluded due to extensive overlap with Google Translate.

The second reliability assessment measures the similarity between the source text and a final translation that we reverse-translate to its original language. This process determines a translation's stability over multiple machine translations and thus the potential robustness of the application. We calculated similarity scores using the online plagiarism checker CopyLeaks at three different levels of precision, from exact matches to translations

involving synonyms, rounded to the nearest whole number. Translations were checked through one (English or German) and two (English then German and German then English) translations before being returned to their original language.

The stability scores show that DeepL generally outperformed Google and Microsoft in these round-robin exercises (DeMattee et al. 2022, 14). DeepL appeared most stable for Chinese, French, and Spanish content and produced average scores at or above 50%; when the translation was reverse-translated, at least half of the words were identical, near, or related matches. Google and Microsoft were most stable for Spanish, French, and Portuguese content. DeepL's strongest outperformance was with the Chinese legal text. The reliability tests suggest that scholars might prefer different MT software for some languages over others. DeepL performs better for translations from Chinese, while Google or Microsoft works better for Romance languages. The general underperformance of machine translators for Japanese is notable.[1]

### Limitations of Machine Translation

While our analysis suggests many benefits of adopting MT applications, we highlight three issues that researchers should consider. First, MT is much easier with newer laws, often published on national websites as HTML or machine-readable PDF, which are quickly and easily read by any application. Older laws may only be available as an image file rather than as text, requiring retyping and careful reformatting, which can induce human error. Second, laws (at least as regards associations and charities) are becoming longer over time, and free MT applications limit file upload size. Most MT applications require legal texts to be divided into smaller individual PDF documents before translation. A potential time-saving solution is to use MT applications to translate the table of contents or index of the law in one PDF, then translate the relevant sections. Finally, it is difficult with current tools to present both original and translated texts in parallel (side-by-side pages), which makes it slightly more challenging to evaluate the quality of the translation by comparing relatable blocks of text.

## Conclusion

The "laws-in-translation problem" exists because governments rarely translate their laws into multiple languages and translating these legal texts typically requires choosing between affordability and precision. While translation professionals provide high-quality translations, these services quickly become cost prohibitive.[2] MT applications are an attractive alternative.

They are increasingly accessible, fast, and affordable. Still, scholars may be reluctant to use these applications for research purposes because the accuracy of the translations is unknown. Moreover, other researchers—specifically journal reviewers—may not be convinced by research findings that depend solely on translation applications. This adds another dimension to the laws-in-translation problem and the challenge of transparently and rigorously translating legal text for comparative research purposes.

We used similar legal text from five countries with diverse languages to assess the performance of three translation applications using five measures: generalizability, flexibility, presentation, simplicity, and reliability. Our assessment found translation applications to be effective but not precise or consistent enough to warrant use without verification. Minor errors aside, these automated tools occasionally make critical errors and lose a text's original meaning. MT tools may thus adversely affect research outcomes. By comparison, humans produce translations without critical errors or deviations from the text's original meaning, although human translators do take issue with other translators' texts.

These findings lead us to recommend that researchers pair machine translations with human translators to produce reliable and affordable translations. Extrapolating from data from our human evaluators, it takes an individual fluent in the necessary languages three times longer to produce a translated text than it does to edit an MT translation of the same source text. This suggests it is far more efficient for a researcher to use *any* translation application to make an initial translation, and then employ a human translator to improve that translation by making minor edits and correcting critical errors as required. As legal texts become longer, more complicated, with more cross-referencing, the efficiency gains from using an MT tool will increase. The use of the recommended protocol also provides a clear and tested method for translation that can be easily explained in future publications, increasing replicability and transparency in comparative law and policy research. We also suggest that researchers who adopt this process deposit and share translations in the public domain. Scholarly repositories (e.g., the Qualitative Data Repository), independent organizations (e.g., the International Center for Not-for-Profit Law), and personal websites can each host translated versions of these public documents.

The hybrid method for translating legal text that we have introduced can expand and improve comparative law and socio-legal research by drastically reducing

---

1 The mechanism causing this underperformance is beyond our scope, we simply note its existence.
2 The price for such services—even the most economical—begins at $25 per page or $0.05 per word.

linguistic knowledge as a constraint. The reduction in the time and translation costs opens new areas of research that would otherwise require large teams and grants to even consider. Accurate translations of new categories of law enables the systematic investigation of the evolution of law over time or the study of diffusion effects both within regions and globally. Generations of political scientists and legal scholars have emphasized the distinction between laws in the books and laws in action. This new method enables scholars to examine the extent to which *de facto* enforcement diverges from *de jure* rules. The large number of languages now available for machine translation, as well as the growth of international research societies providing networks with native speakers, enables research to redress the relative marginalization of some geographic areas that are either too poor to translate and publish their laws, or where legal text is written in a language that is unfamiliar to comparative legal scholars. Finally, the output of MT applications works easily as raw data for qualitative research software such as NVivo and ATLAS.ti.

## References

Bennett, Andrew. 2007. "Symposium: Multi-Method Work, Dispatches from the Front Lines." *Qualitative Methods* 5, no. 1 (Spring): 9-28.

Berinzon, Maya, and Ryan C. Briggs. 2019. "Measuring and Explaining Formal Institutional Persistence in French West Africa." *The Journal of Modern African Studies* 57, no. 2 (June): 183-202. https://doi.org/10.1017/S0022278X19000077.

Bloodgood, Elizabeth A., Joannie Tremblay-Boire, and Aseem Prakash. 2014. "National Styles of NGO Regulation." *Nonprofit and Voluntary Sector Quarterly* 43, no. 4 (August): 716-36. https://doi.org/10.1177/0899764013481111.

Bradford, Anu, Adam S. Chilton, Christopher Megaw, and Nathaniel Sokol. 2019. "Competition Law Gone Global: Introducing the Comparative Competition Law and Enforcement Datasets." *Journal of Empirical Legal Studies* 16, no. 2 (June): 411-43. https://doi.org/10.1111/jels.12215.

Cook-Martín, David, and David Scott FitzGerald. 2019. "How Their Laws Affect our Laws: Mechanisms of Immigration Policy Diffusion in the Americas, 1790–2010." *Law & Society Review* 53, no. 1 (March): 41-76. https://doi.org/10.1111/lasr.12394.

DeMattee, Anthony J. 2020. "Domesticating Civil Society: How and Why Governments Use Laws to Regulate CSOs," PhD diss., Indiana University.

DeMattee, Anthony J. 2022a. "Are Constitutions Important in Diffusion Research?: How Constitutional-Level Rules, Autocratic Learning, and Emulation Shape the Development of Civil Society Laws in East Africa." Presented at the Workshop on Openness in Law-Related Research, New York City, New York, April.

DeMattee, Anthony J. 2022b. "A Grammar of Institutions for Complex Legal Topics: Resolving Statutory Multiplicity and Scaling-up to Jurisdiction-Level Legal Institutions." *Policy Studies Journal* (Forthcoming):1-34.

DeMattee, Anthony J., Nick Gertler, Takumi Shibaike, and Elizabeth A. Bloodgood. 2022. "Supplemental Information For: Overcoming the Laws-in-translation Problem: Comparing Techniques to Translate Legal Texts." OSF Preprints. February 23. doi:10.31219/osf.io/jc5p9.

Diakopoulos, Nicholas. 2013. "Algorithmic Accountability Reporting: On the Investigation of Black Boxes." *Tow Center for Digital Journalism, Columbia Journalism School.* https://doi.org/10.7916/D8ZK5TW2.

Elkins, Zachary, and Tom Ginsburg. 2021. "Characteristics of National Constitutions, Version 3.0." *Comparative Constitutions Project.* https://comparativeconstitutionsproject.org/download-data/.

Elkins, Zachary, Tom Ginsburg, and James Melton. 2009. *The Endurance of National Constitutions.* Cambridge: Cambridge University Press.

Elman, Colin, Diana Kapiszewski, and Arthur Lupia. 2018. "Transparent Social Inquiry: Implications for Political Science." *Annual Review of Political Science* 21, no. 1 (May): 29-47. https://doi.org/10.1146/annurev-polisci-091515-025429.

Giampieri, Patrizia. 2016. "A Critical Comparative Analysis of Online Tools for Legal Translations." *The Italian Law Journal* 2 (2): 445-61.

Glasius, Marlies, Jelmer Schalk, and Meta De Lange. 2020. "Illiberal Norm Diffusion: How Do Governments Learn to Restrict Nongovernmental Organizations?" *International Studies Quarterly* 64, no. 2 (June): 453-68. https://doi.org/10.1093/isq/sqaa019.

Goemans, Hein E. 2007. "Qualitative Methods as an Essential Complement to Quantitative Methods." *Qualitative Methods* 5, no. 1 (Spring): 11-13.

Hummel, Calla, John Gerring, and Thomas Burt. 2021. "Do Political Finance Reforms Reduce Corruption?" *British Journal of Political Science* 51, no. 2 (April): 869-89. https://doi.org/10.1017/S0007123419000358.

Kapiszewski, Diana, and Sebastian Karcher. 2021. "Transparency in Practice in Qualitative Research." *PS: Political Science & Politics* 54, no. 2 (April): 285-91. https://doi.org/10.1017/S1049096520000955.

Law, David S., and Mila Versteeg. 2013. "Sham Constitutions." *California Law Review* 101, no. 4 (August): 863-952.

Lemon, Edward, and Oleg Antonov. 2020. "Authoritarian Legal Harmonization in the post-Soviet Space." *Democratization* 27 (7): 1221-39. https://doi.org/10.1080/13510347.2020.1778671.

Lin, Yu-Hsin, and Yun-chien Chang. 2018. "An Empirical Study of Corporate Default Rules and Menus in China, Hong Kong, and Taiwan." *Journal of Empirical Legal Studies* 15, no. 4 (December): 875-915. https://doi.org/10.1111/jels.12200.

Lucas, Christopher, Richard A. Nielsen, Margaret E. Roberts, Brandon M. Stewart, Alex Storer, and Dustin Tingley. 2015. "Computer-Assisted Text Analysis for Comparative Politics." *Political Analysis* 23, no. 2 (Spring): 254-77. https://doi.org/10.1093/pan/mpu019.

Pound, Roscoe. 1910. "Law in Books and Law in Action." *American Law Review* 44, no. 1 (January-February): 12-36.

Salminen, Joni, Maximilian Hopf, Shammur A. Chowdhury, Soon-gyo Jung, Hind Almerekhi, and Bernard J. Jansen. 2020. "Developing an Online Hate Classifier for Multiple Social Media Platforms." *Human-centric Computing and Information Sciences* 10 (1): 1-34. https://doi.org/10.1186/s13673-019-0205-6.

# QMMR
Qualitative &
Multi-Method
Research

# Symposium

# Back to the Field: Uncertainty and Risk in Field Research

Jannis J. Grimm
*Freie Universität Berlin*

Sarah E. Parkinson
*Johns Hopkins University*

Ellen Lust
*University of Gothenburg*

Isabell Schierenbeck
*University of Gothenburg*

Kevin Koehler
*Sant'Anna School of Advanced Studies - Pisa*

Dina Zayed
*University of Sussex*

The rapid spread of COVID-19 beginning in early 2020 caused global disruption. As the risk of infection rose and public health authorities around the world enacted measures to contain the virus, everyday life ground to a halt. Activities that seemed routine in late 2019 became fraught with uncertainty. Fieldwork was no exception. Most field researchers had to change or cancel at least some of their plans; some left their field in a hurry before travel was shut down while others had to lock down on site; most academic institutions restricted travel, with some even prohibiting all forms of international movement. In brief, many traditional forms of fieldwork became all but impossible during the pandemic.

Even as parts of the world begin to emerge from the pandemic, things have not returned to normal. Indeed, the emergence of the Omicron variant in November 2021 led to new restrictions, with some universities again moving to block field research.[1] It is important to note that such restrictions sometimes seem to be driven by factors other than the risk of infection alone. Infection rates in parts of Europe or the US were frequently just as high or even higher than in other parts of the world, yet restrictions seemed to be primarily aimed at preventing movement between the Global North and South. At the same time, global vaccine inequalities and vaccine resistance are threatening to relegate parts of the world to the category of places which are not safe

---

1 For example, Arantxa Rodriguez-Uribe (@MAranzazuRU, November 30, 2021) tweeted that Princeton had just banned all international fieldwork.

for research. This is a worrying trend that might have serious consequences for knowledge production, as research efforts are being shifted to more "accessible" fields. It also further exacerbates existing inequalities in the discipline.

In this symposium, we reflect on lessons we can draw for fieldwork safety. These reflections build on our work on *Safer Field Research in the Social Sciences* (Grimm et al. 2020), a handbook to which most of the authors in this symposium contributed. The pandemic has created an unprecedented level of awareness of fieldwork risk. While attention to the issue is welcome—in fact, many of us had called for more awareness before the pandemic (Grimm et al. 2020; Lake and Parkinson 2017)—the way many academic institutions have reacted to the crisis has highlighted core weaknesses in how we think of and approach fieldwork risk.

To begin with, uncertainty in field research is not new. In fact, field research has always been subject to uncertainty and risk, especially for those who work in environments characterized by conflict or political repression (Bond, Lake, and Parkinson 2020; Glasius et al. 2018; Grimm et al. 2020; Mac Ginty, Brett, and Vogel 2020). The pandemic has certainly brought such challenges into even starker relief, yet COVID-19 is far from being the only source of risk. What is more, compared to other types of dangers, the threat of infection can be managed through vaccination, mask wearing, hygiene rules, and regular, accessible testing. While many academic institutions have implemented such measures to maintain in-person teaching despite high infection rates, mitigation measures are often not considered sufficient to enable safe fieldwork even where they are available.

We advocate for a shift to a risk management perspective. Many forms of fieldwork are risky, and they were risky before the pandemic as well. Instead of restricting physical access to the field, academic institutions should facilitate the clear-eyed management of these risks. In fact, those of us who managed to continue fieldwork during the pandemic can contribute valuable lessons in this regard (Lust and Schierenbeck, this symposium). Such a shift would require addressing perverse incentives in risk assessment procedures which frequently appear as administrative hurdles to researchers (see Koehler, this symposium); it would also necessitate a level-headed look at the risks associated with online forms of data collection which are frequently touted as alternatives (Grimm, this symposium). Finally, it would imply a different culture of academic advising that addresses the ethical and safety challenges of fieldwork (Parkinson and Zayed, this symposium). We highlight these issues in the hope that we can contribute to a

discussion on the future of fieldwork as we slowly begin to move back to a (new) normal.

## Back to the Field

Early discussions on fieldwork during Covid understandably focused on how researchers could adapt (Lupton 2021). Initial measures included the increasing use of virtual platforms for data collection, such as through online interviews (Howlett 2022; Vokes and Atukunda 2021), phone interviews, or online panels instead of face-to-face survey research (Arechar and Rand 2021; Will, Becker, and Weigand 2020), or online recruitment for field experiments (Li et al. 2021), for example. Others capitalized on the fact that political activity also moved online during the pandemic. This meant that some projects could explore virtual fields, collecting data directly on Twitter, Facebook, Telegram or other online spaces (Christia and Lawson 2020; Käihkö 2020). Still others increased their reliance on local collaborators who could still enter the field (Kamara, Mokuwa, and Richards 2020).

Some of these innovations were pandemic-era stop-gap measures. Others reflect larger trends that have accelerated during the pandemic, or long-standing issues highlighted by the coronavirus crisis. Using remote techniques for qualitative data collection, for example, is certainly no invention of the last two years. There is a large literature on online research methods, including on the advantages and drawbacks of such approaches (Fielding, Lee, and Blank 2008; Namey et al. 2020). But virtual forms of fieldwork have become much more prominent during the pandemic. Similarly, the ethics and logistics of working with research assistants have also been discussed before the pandemic (Cronin-Furman and Lake 2018; Leck 2014), even though the COVID-19 crisis has given increased urgency to these debates (Nyenyezi Bisoka 2020; Rudling 2021). From a research ethics perspective, increased reliance on research assistants becomes problematic if it is seen as a risk avoidance strategy for researchers unable or unwilling to travel. It is not; it merely shifts risks from researchers to their interlocutors or research assistants (see Grimm, this symposium).

The pandemic has highlighted the uncertainty associated with fieldwork. While this is old news for those of us working in contexts of political conflict and repression, the COVID-19 crisis has created an unprecedented degree of attention to issues of fieldwork safety. We argue that we should take this opportunity to address the uncertainty associated with field research and to review some of the processes we routinely implement. As Ellen Lust and Isabell Schierenbeck suggest in their essay, the COVID-19 pandemic has "fostered practices

that can serve field researchers well. As the pandemic subsides and fieldwork resumes, we should make sure that these practices are kept" (Lust and Schierenbeck, this symposium).

Which pandemic innovations and practices are likely to be of continued relevance in post-pandemic fieldwork? In their opening essay, Ellen Lust and Isabell Schierenbeck identify five broad lessons. They argue that the pandemic has (again) highlighted the importance of risk assessment, the pervasiveness of threats, the crucial role of adequate and up-to-date information, the need to recognize and take responsibility for ways in which our research might endanger others, as well as the issue of mental health during and after fieldwork. While none of these issues are entirely new, the pandemic highlighted their importance for researchers who were not used to thinking about their work in terms of uncertainty and risk.

Recognizing uncertainty also means developing ways of managing it. Unfortunately, many universities and research institutions have reacted to the pandemic by restricting research rather than enabling safer practices. Kevin Koehler argues that this has highlighted structural features in the way in which we conduct risk assessments. Institutional risk assessment procedures create perverse incentives for researchers. Since risk assessments frequently determine access to funding, researchers face incentives to downplay risk so as not to jeopardize their fieldwork. At the same time, the pandemic has demonstrated that such forms of risk assessment are not particularly helpful in actual crisis situations as they do not lead to appropriate contingency planning. Rather than threatening to restrict research, risk assessment procedures should be occasions for information exchange and learning which enable safer research.

The tools we increasingly resort to when we conduct online fieldwork are at the center of the contribution by Jannis Grimm. Given that researchers' own understanding of these tools is frequently limited, their uncritical use might lead to the outsourcing of risk to interlocutors. While scholars may conduct their research from the safety of their own homes, their interlocutors are left to worry about the potential of (online) surveillance, safe and sufficient internet access, and data security. At the same time, online research methods can create "affective detachment," not only when it comes to the potential risks associated with the research process itself, but also in terms of the everyday needs of their interlocutors and the real-world problems they face. As the pandemic continues to impede traditional field research in many parts of the world, academic institutions can help their researchers in navigating the ethical dilemmas of remote research by establishing workflows and support structures that specifically address the risks associated with increasing technological dependence.

Finally, Sarah Parkinson and Dina Zayed introduce the notion of "reflexive advising" as a tool for managing uncertainty. Noting that academic advisors tend to significantly shape their mentees' research projects yet are often absent from discussions on risk and uncertainty, they emphasize the need for advisers and mentees to "actively and collectively evaluate a combination of researcher positionality and contextual factors in order to open discussions of field safety" (Parkinson and Zayed, this symposium). Advisers also need to be aware of their own limitations and should actively support their mentees in seeking the feedback of relevant disciplinary networks. Such new forms of advising could go a long way in creating awareness of risks beyond COVID-19 while enabling ethically sound risk management practices.

## Conclusion

As we consider ways of "returning to normal" in the fieldwork-based social sciences, scholars should heed lessons learned during the pandemic. The disruptions of COVID-19 have highlighted core weaknesses in institutional responses to fieldwork risk, as well as worrying trends of restricting research and outsourcing risks. The specific risks associated with the pandemic are certainly real, yet mitigation strategies are well known and are becoming increasingly available to researchers. At the same time, some of the solutions implemented to keep fieldwork running during the pandemic have worrisome ethical implications. It is time to take fieldwork risk for what it is—a set of challenges to be recognized and managed— and not as a collection of problems to be avoided or outsourced. We hope that the experience of conducting fieldwork during the pandemic will help push such a shift in perspective.

## References

Bond, Kanisha D., Milli Lake, and Sarah E. Parkinson. 2020. "Lessons from Conflict Studies on Research during the Coronavirus Pandemic." *Social Science Research Council: Items*. July 2, 2020. https://items.ssrc.org/covid-19-and-the-social-sciences/social-research-and-insecurity/lessons-from-conflict-studies-on-research-during-the-coronavirus-pandemic/.

Christia, Fotini, and Chappell Lawson. 2020. "Training the Covid-19 Cohort: Adapting and Preserving Social Science Research." *Social Science Research Council: Items*. July 30, 2020. https://items.ssrc.org/covid-19-and-the-social-sciences/social-research-and-insecurity/training-the-covid-19-cohort-adapting-and-preserving-social-science-research/.

Cronin-Furman, Kate, and Milli Lake. 2018. "Ethics Abroad: Fieldwork in Fragile and Violent Contexts." *PS: Political Science & Politics* 51, no. 3 (July): 607–14. https://doi.org/10.1017/S1049096518000379

Fielding, Nigel, Raymond M Lee, and Blank, Grant. 2008. *The SAGE Handbook of Online Research Methods*. Thousand Oaks: SAGE Publishing. https://dx.doi.org/10.4135/9780857020055.

Glasius, Marlies, Meta de Lange, Jos Bartman, Emanuela Dalmasso, Aofei Lv, Adele Del Sordi, Marcus Michaelsen, and Kris Ruijgrok. 2018. *Research, Ethics and Risk in the Authoritarian Field*. Basingstoke: Palgrave Macmillan. https://doi.org/10.1007/978-3-319-68966-1.

Grimm, Jannis, Kevin Koehler, Ellen 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 Publishing.

Howlett, Marnie. 2022. "Looking at the "Field" through a Zoom Lens: Methodological Reflections on Conducting Online Research during a Global Pandemic." *Qualitative Research* 22, no. 3 (June): 387-402. https://doi.org/10.1177/1468794120985691.

Käihkö, Ilmari. 2020. "Conflict Chatnography: Instant Messaging Apps, Social Media and Conflict Ethnography in Ukraine." *Ethnography* 21, no. 1 (March): 71–91. https://doi.org/10.1177/1466138118781640.

Kamara, Foday Mamoud, Esther Yei Mokuwa, and Paul Richards. 2020. "How Villagers in Central Sierra Leone Understand Infection Risks under Threat of Covid-19." *PLOS ONE* 15 (6): e0235108. https://doi.org/10.1371/journal.pone.0235108.

Lake, Milli, and Sarah E. Parkinson. 2017. "The Ethics of Fieldwork Preparedness." *Political Violence @ a Glance*. June 5, 2017. https://politicalviolenceataglance.org/2017/06/05/the-ethics-of-fieldwork-preparedness/.

Leck, Hayley. 2014. "Whose Voice? Ethics and Dynamics of Working with Interpreters and Research Assistants." In *Fieldwork in the Global South*, edited by Jenny Lunn, 59-68. London: Routledge.

Li, Zhi, Po-Hsuan Lin, Si-Yuan Kong, Dongwu Wang, and John Duffy. 2021. "Conducting Large, Repeated, Multi-Game Economic Experiments Using Mobile Platforms." *PLOS ONE* 16 (4): e0250668. https://doi.org/10.1371/journal.pone.0250668.

Lupton, Deborah. 2021. "Doing Fieldwork in a Pandemic." Crowdsourced Document. https://docs.google.com/document/d/1clGjGABB2h2qbduTgfqribHmog9B6P0NvMgVuiHZCl8/edit#heading=h.ze8ug1cqk5lo.

Mac Ginty, Roger, Roderick Leslie Brett, and Birte Vogel, eds. 2020. *The Companion to Peace and Conflict Fieldwork*. Basignstoke: Palgrave Macmillan.

Namey, Emily, Greg Guest, Amy O'Regan, Christine L. Godwin, Jamilah Taylor, and Andres Martinez. 2020. "How Does Mode of Qualitative Data Collection Affect Data and Cost? Findings from a Quasi-Experimental Study." *Field Methods* 32, no. 1 (February): 58–74. https://doi.org/10.1177/1525822X19886839.

Nyenyezi Bisoka, Aymar. 2020. "Disturbing the Aesthetics of Power: Why Covid-19 Is Not an "Event" for Fieldwork-Based Social Scientists." *Social Science Research Council: Items*. May 28, 2020. https://items.ssrc.org/covid-19-and-the-social-sciences/social-research-and-insecurity/disturbing-the-aesthetics-of-power-why-covid-19-is-not-an-event-for-fieldwork-based-social-scientists/.

Rodriguez-Uribe, Arantxa. 2021. "Princeton just banned all international fieldwork due to Omicron." Twitter, November 30, 2021. https://twitter.com/MAranzazuRU/status/1465719388229685251

Rudling, Adriana. 2021. "Now Is the Time to Reassess Fieldwork-Based Research." *Nature Human Behaviour* 5, no. 8 (August): 967. https://doi.org/10.1038/s41562-021-01157-x.

Vokes, Richard, and Gertrude Atukunda. 2021. "Fieldwork through the Zoomiverse: Sensing Uganda in a Time of Immobility." *Anthropology in Action* 28, no. 1 (March): 73–8. https://doi.org/10.3167/aia.2021.280114.

Will, Gisela, Regina Becker, and Dominik Weigand. 2020. "COVID-19 Lockdown during Field Work." *Survey Research Methods* 14 (2): 247-52. https://doi.org/10.18148/SRM/2020.V14I2.7753.

# Lessons for Safer Fieldwork Practices from the Covid Pandemic

Ellen Lust
*University of Gothenburg*

Isabell Schierenbeck
*University of Gothenburg*

Pandemics can provide important lessons and foster better practices. The Bubonic Plague in the 1300s led to the establishment of quarantines for patients with contagious diseases (Pearce Rotondi 2020); a Tuberculosis outbreak in the early 20th century led to the expansion of open-air schools (Blei 2020) and the Spanish Flu of 1918 pushed states to develop public health programs (and fostered the take-off of the paper dixie cup) (Spinney 2017).

Lessons for fieldwork may be less obvious, but they are no less important. The pandemic has highlighted the uncertainty surrounding research risks. As we discuss in *Safer Field Research in the Social Sciences* (Grimm et al. 2020, 6), fieldwork entails both "the probability that some threat to the research project is realized, and the severity of the impact this threat has on the project." Often, as with the pandemic, there is uncertainty around both the probability and severity of threats, and thus the implications of choices for researchers, their partners, and others. We argue that the pandemic has highlighted the problems of uncertainty and fostered practices that can serve field researchers well. As the pandemic subsides and fieldwork resumes, we should make sure that these practices are kept.

## Lesson 1: The Importance of Ongoing Risk Assessments

The pandemic has fostered a practice of on-going risk assessments. Before the pandemic, vacation plans were made, wedding and birthday celebrations prepared, and conferences set months in advance, with little consideration of circumstance as the event approached. After more than two years of COVID-19, we have seen plans made and unmade, new opportunities open and others close. We have all become keenly aware that what may be a reasonable plan today may be untenable tomorrow, as the coronavirus variant, vaccine prevalence, travel restrictions or other conditions change. And we have learned to undertake on-going assessments.

The same is true with fieldwork. The traditional approach to fieldwork is to do risk assessments before fieldwork begins. As researchers prepare for a project, they consider risks, file for ethics reviews, receive IRB approval, and begin fieldwork, never looking back. Researchers are aware of risks and engage in assessments—taking into account the likelihood that a threat may materialize, and the costs of it when it occurs—but they assume risks are constant.

Yet, conditions can and do change. Well after plans are made, approvals given, and research has begun, political tensions escalate, social backlash to the research mounts, climate disasters strike, and other health crises break out. At times these are obvious: military coups, popular demonstrations, and hurricane force winds are all easy to recognize. More often, they are more localized and subtle, at least before disaster strikes. Researchers may not readily spot increased state surveillance or neighbors' suspicions, particularly if they assume the research environment poses no risks.

Thus, researchers need to shift from a perspective that views risk analyses as a step in the research process, undertaken before fieldwork begins, to one that understands the need for ongoing risk analysis. In our book, we argued for the importance of constant risk assessments, beginning before research, and continuing throughout the process. Researchers recognize shifting threats only if they are attuned to the possibility that new threats may emerge, or their potential to materialize increases.

Our own research experience during the pandemic brought such considerations to the fore. Early in the pandemic, researchers at the Program on Governance and Local Development (GLD) sought to draw on previous research in Malawi, implementing surveys aimed at understanding the challenges that communities face and providing valuable insights to key stakeholders. We believed the project would be most effective if we could gather telephone numbers from communities where we previously ran our surveys but knew that doing so would require returning to these communities. We weighed the risks of sending researchers to communities, albeit early in the pandemic, with the benefits of providing information as conditions worsened. We developed a procedure that armed researchers with personal protective equipment (PPE) and minimized their contact with community members (they met one community member who gathered the information), and we proceeded to implement the project. The experience not only emphasized the importance of ongoing risk

assessments, it also demonstrated that these assessments could lead to the adjustment of plans, and project abandonment may not be necessary. This is not new; yet the ubiquitous nature of the pandemic highlighted the ever-present importance of ongoing risk assessments.

## Lesson 2: The Pervasiveness of Threats

The COVID-19 pandemic has also drawn our attention to the pervasiveness of threats. The threat of coronavirus was not limited to public or private spaces, outdoors or inside. Suddenly, going to restaurants or grocery stores, workplaces or apartments all had an element of risk. The level of risk varied, but every encounter required thought.

The same holds for fieldwork. Fieldwork risks extend beyond the walls of the office or training rooms. As we discuss in Chapter 3 of *Safer Field Research in the Social Sciences* (Grimm et al. 2020), the basic details of life (e.g., finding an apartment, setting up an office, hiring assistants, navigating transportation) entail potential risks and unknowns. This is particularly true for those moving to a new field site. Whether this is a new neighborhood in one's hometown or across the ocean, researchers need to assess the potential threats. Considering the intersection of specific research and multipliers of risk (e.g., political and social structures, methods of data collection) leads to more effective and comprehensive risk assessments.

Consider, for example, survey research. Before the pandemic, one could easily implement household surveys, sending teams into the field to interview individuals in houses, stay overnight in hotels, eat at restaurants, and eventually return to their homes. After the pandemic, not only were interviews potentially problematic, hotels, restaurants, and cross-country travel all risked spreading disease. This led our and other teams to re-evaluate the mode of implementation, turning from face-to-face to telephone surveys, despite their drawbacks.

## Lesson 3: The Importance of Information

The pandemic has also highlighted the importance of information. At first, little was known about the virus. Was it spread through air or water? Could it be carried on surfaces? What places and activities were safe, and which were not? How dangerous was it, really? And who was best placed to answer these questions? From the very start of the pandemic, there was a lack of coherent and reliable information, and recommendations and strategies became both highly politicized and globally diverse. Thus, the answers to these questions were not always obvious. The challenge was to be cautious but calm, to gather as much information as possible about safe practices, and to do so from credible sources.

Field researchers need to take a similar approach.

Researchers need to gather as much information as possible, both before and during fieldwork. The pandemic brought this lesson to the fore. People became well-practiced in checking government websites for advice on travel safety before boarding planes. Researchers need to do the same. Moreover, because fieldwork threats are all-encompassing, this information gathering should entail mundane decisions as well. When finding housing, office sites, and schools for children, researchers need to consider not only quality, locations, and prices, but also potential security risks.

Researchers also need to examine their own vulnerabilities and capabilities as they gather information and assess risks. During the pandemic, individuals understood that the risks entailed in different actions depend on multiple factors, such as age, underlying health conditions, and vaccine status. The extent to which a neighborhood or form of transportation is risky similarly depends not only on the place but also on the person. It may be very different for a tough-looking, burly man to undertake ethnographic work while living on the streets of New York City than it would be for a petite woman. It is not the same for a freshly minted MA student, with limited language capabilities, to engage in research in a conflict-ridden environment as it is for a seasoned researcher with advanced language skills. Researchers need to consider their ability to understand the context, the extent to which they convey themselves to be outsiders (due to ethnicity, class, or other factors), and other characteristics that may impact whether it is safe to engage in the fieldwork.

Finally, researchers, as well as supervisors and donors, need to reflect on how much uncertainty they are willing to accept. There is no clear rule regarding acceptable uncertainty. As we have seen during the pandemic, very reasonable people can disagree about what "reasonable risks" entail. Yet, the research community should have open and ongoing conversations about how to assess and weigh uncertainty and risk, keeping in mind the costs borne both by researchers and others.

## Lesson 4: You as Potential Threat; Taking Responsibility

The pandemic has helped us recognize that we are not only the potential subject of threats, but carriers as well. We can transmit COVID-19 as much as we can catch it. Consequently, most of us have come to accept that we have an obligation to vaccinate and wear masks, not only to shield ourselves, but also to protect others.

This was particularly evident as we conducted fieldwork under pandemic conditions. In the GLD project described above, we recognized the importance of protecting both our researchers and the community

members they met. Thus, we provided PPE to researchers, but also instructed them to minimize engagement with community members and to take social distancing seriously. Moreover, apart from the first canvassing to gather phone numbers (conducted early in the pandemic, as only the first cases reached Malawi), we trained researchers remotely and conducted all interviews via telephone. As researchers, we recognized that we posed threats to the interviewers, and that they in turn were potential threats to the interviewees.

Even in non-pandemic situations, researchers need to recognize that they are not only under threat but can pose threats for others. For instance, when foreigners engage in fieldwork, their very presence may create problems for friends, assistants, and others associated with them. So, too, when researchers require associates to travel to remote areas or on dangerous roads to reach them, they put others at risk.

Researchers thus have an obligation to implement research in ways that do not put others in compromising positions. This is particularly true given the power dynamics often at play in these settings. Associates are often unwilling to voice concerns when working with researchers who are seen as more powerful, educated, or the source of livelihoods. It is up to the researcher, then, to consider the safety of others as carefully as they do their own.

## Lesson 5: Minding Mental Health

Finally, the pandemic has also taught us the importance of minding mental health. The pandemic resulted in new forms of stress and trauma. The very uncertainty of the pandemic created stress. First, the lack of information, described above, meant individuals were never quite sure of the risks, or when the pandemic would end. Second, the pandemic has been heavy with traumatic experiences. Witnessing people dying from COVID-19, inadequate healthcare systems, unequal distribution of vaccines, and the effects of lockdowns caused stress for everyone. Third, the pandemic made everyone a source of threat. It became increasingly clear to many that they were not only at risk, they could also put others at risk. And finally, the pandemic created new professional obstacles and fostered a great deal of economic insecurity. As Kevin Koehler notes in this symposium, these issues have been worse for some than others—with junior scholars and women facing challenges disproportionately. But for many, the challenges of managing fieldwork and on-line teaching, along with children, elderly parents, and an ever-shifting landscape of pandemic restrictions, has created unprecedented mental strain.

Consequently, the pandemic highlighted the importance of mental self-care. Deteriorating mental health leads individuals to make poor choices, exacerbating problems, and putting others further at risk. Ongoing attention to mental health is critical and needs to be given during periods of stress and uncertainty, and not made to wait until the pandemic is over.

Fieldwork oftentimes entails similar sets of stress and requires the same attention to mental well-being (Grimm et al. 2020, 80-81). There is often uncertainty, a lack of information, a potential for traumatic experiences, and concern that a researcher is not only at risk but can put others at risk as well. Researchers engaged in fieldwork should pay attention to their mental health, the well-being of those around them, and reach out for support. So too should the PhD supervisors, colleagues, and family members who support researchers during fieldwork. Doing so not only supports the researcher, but also helps to ensure clear-eyed choices and reduces risks to those around them.

## Practical Solutions

Two practical solutions can help researchers address the lessons learned.

### Maintaining Networks

Fieldwork requires researchers to develop and maintain a network of trusted colleagues, both at the field site and away. Those at the site are often best situated to help answer questions, provide feedback, and assist when problems arise. To locate a "field partner" to help monitor one's exposure to changing field conditions can be valuable. The "field partner" can be a colleague, supervisor, friend, or someone who knows the challenges of fieldwork. It is someone with whom you can have frequent (weekly or monthly) check-ins, who can listen, reflect, and mirror reflections, emotions, and issues raised in the conversations. Friends, family, and associates away from the field site are also important. They are often well-placed to notice a change in tone or demeanor that can reflect increasing stress, or to notice changes in circumstances that those closer to the situation may miss.

Moreover, diversity within the network is important, as those from different backgrounds and positions are better placed to receive new information or offer a different perspective. Researchers need to consider from whom they get credible information to answer these questions. Assessing fieldwork risks requires one to understand the research question and method, fieldwork context, and researchers' experience and qualities. It also requires multiple perspectives from trusted individuals. Networks are particularly helpful in locating individuals who can answer these questions frankly. However, researchers should seek out multiple points of view,

from diverse perspectives, to best assess security risks.

It takes time to build a strong network of trusted individuals, with strong ties based on respect and an understanding of shared values, identities, and consequences. Thus, researchers should begin establishing networks long before fieldwork begins, and seek to extend their networks throughout their fieldwork. They should do so cautiously, as unfortunately, not all contacts are benign.

But they should also do so consciously and consistently. A recent study on the effect of the pandemic on networks found that while exchanges among close associates increased, those with more distant associates shrank (King and Kovács 2021). In many ways, the stress and isolation of the pandemic is similar to the stress and isolation of prolonged fieldwork in difficult settings. The challenge, then, is to resist "turtling in"—or losing outer layers of networks.

### Mindfulness

From previous studies on secondary trauma, we know that preparation is key in prevention and a vital part of self-care (Močnik 2020). First, the researcher needs to have strategies for how to recognize their own vulnerabilities and potential triggers. Knowing yourself and being aware of your baselines is important. One suggestion is to undertake pre-fieldwork reflection, recording potential causes for distress and coping strategies should it arise. Here, one can reflect on past traumatic events and experiences to consider what types of encounters trigger stress and understand how one reacts. One can also talk to mentors and colleagues to hear their experiences and mentally prepare for problems they may encounter.

Second, the researcher should practice self-care before, during, and after fieldwork. This could include everyday self-care and coping strategies, such as good nutrition, regular exercise, mediation, creativity, and spirituality, but it is important to consider how self-care can be exercised in different and changing settings. Self-care should also include establishing routines. These need to be planned and thought through well before fieldwork, although uncertainty should be taken into consideration. One might want to plan for physical exercise, and make sure to prepare yourself to keep in contact with people (at home and in the field) in various ways. If mobility is restricted, as during the pandemic, and one is forced into isolation or stuck in the field without a possibility to exit, it is vital to have prepared for how to keep mentally sane.

Third, maintaining a diary is a useful strategy for forcing oneself to set aside time daily to reflect on broader surroundings, engagements with others, and changing circumstances. These systematic and on-going reflections on fieldwork can help researchers remain aware of emerging risks and react accordingly. This not only prompts the researcher to take time to reflect, but also provides a record of change, and can help make well-balanced judgements and decisions for oneself, research partners, and others. Reviewing past notes can also help highlight changes and draw attention to emerging risks.

## Conclusion

The pandemic has brought home several lessons that are critical for establishing safer research practices. There is a great deal of uncertainty about the likelihood and significance of potential threats. Risks can change over time, often with little warning. Researchers are not only at risk, but also pose potential threats to those around them. It is thus critical to gather information, update risk assessments, and engage in practices that preserve physical safety and mental health. Doing so not only improves researchers' abilities to implement studies during the pandemic, these practices will be critical to implementation long after the immediate stress of the pandemic has subsided.

## References

Blei, Daniela. "When Tuberculosis Struck the World, Schools Went Outside." *Smithsonian Magazine*. September 1, 2020. https://www.smithsonianmag.com/history/history-outdoor-schooling-180975696/

Grimm, Jannis J, Kevin Koehler, Ellen Lust, Ilyas Saliba, and Isabell Schierenbeck. 2020. *Safer Field Research in the Social Sciences*. Thousand Oaks: Sage Publications.

King, Melissa, and Balázs Kovács. 2021. "Research: We're Losing Touch with Our Networks." Harvard Business Review, February 12, 2021. https://hbr.org/2021/02/research-were-losing-touch-with-our-networks.

Močnik, Nena. 2020. "Re-Thinking Exposure to Trauma and Self-Care in Fieldwork-Based Social Research: Introduction to the Special Issue." *Social Epistemology* 34 (1): 1-11. https://doi.org/10.1080/02691728.2019.1681559.

Pearce Rotondi, Jessica. 2020. "5 Hard-Earned Lessons from Pandemics of the Past." *HISTORY*, August 10, 2020. https://www.history.com/news/pandemics-lessons.

Spinney, Laura. 2017. "How the 1918 Flu Pandemic Revolutionized Public Health." *Smithsonian Magazine*, September 27, 2017. https://www.smithsonianmag.com/history/how-1918-flu-pandemic-revolutionized-public-health-180965025/.

# Managing Uncertainty:
# How Risk Assessment Can Empower Field Research

Kevin Koehler
*Sant'Anna School of Advanced Studies - Pisa*

The discussion on risk in social science fieldwork has seen two diametrically opposed positions. On the one hand, there are those who advocate risk avoidance. Frequently driven by concerns about legal exposure and insurance coverage, such perspectives tend to predominate in university administrations and among funding bodies. They consequently also shape the risk assessment procedures many research projects must clear. Their main aim is to minimize risk by avoiding risky research topics and field sites altogether. Taken to its logical conclusion, however, such a position would endanger knowledge production on issues of central social and political importance.

On the other hand, there are those who not only accept risk, but positively revel in the "edginess" of research conducted in challenging environments. Most researchers active in fieldwork have probably felt, and maybe even occasionally succumbed to, the temptation of recounting the dangerous situations they had to navigate during their fieldwork. Some of us might occasionally even brag about such experiences over drinks at the margins of academic conferences, highlighting our personal resilience and ability to withstand stress. While I personally have learned a lot from listening to the experiences of colleagues, there is a narrow line between exchanging information and an unhealthy dynamic of "out-dangering" (Lake and Parkinson 2017).

These portrayals are, of course, caricatures. Few would probably advocate completely abandoning field research on conflict, organized crime, political violence, or authoritarianism. These issues are simply too important to be ignored. No responsible researcher, in turn, would dispute that there are risks which are simply not worth taking. In between these extremes, however, there is a wide variety of ways in which researchers and research institutions deal with risk and uncertainty. The COVID-19 pandemic has highlighted issues of uncertainty and risk in fieldwork. It has also demonstrated that existing approaches to risk assessment are not particularly useful tools when it comes to addressing acute crises.

In this article, I reflect on the way we conduct risk assessments when designing research projects. I do so in three steps. First, I outline how the pandemic has revealed limitations in our risk assessment procedures. Against

this backdrop, I then suggest how risk assessment can be used to empower, rather than restrict, researchers. Third, I conclude by discussing how trends in field research practices which have been accelerated by the pandemic, in particular the outsourcing of fieldwork and the digitalization of field research, have exacerbated specific forms of risk.

## Fieldwork Was Always Uncertain

The pandemic and the various containment measures taken by different governments have made many "traditional" forms of fieldwork impossible and have thus disrupted ongoing fieldwork processes. Many researchers found that they had to leave their field sites early due to the evolving epidemiological situation (Fikrig 2020), or that they could not implement projects as planned and had to redesign them as they went along (MacLean et al. 2021). This has led to fruitful debates on the nature of fieldwork during a global pandemic, the possibility of conducting (some) forms of fieldwork online, as well as the prospects of "returning to normal" (Baczko and Dorronsoro 2020; Krause et al. 2021; Wood et al. 2020).

While this attention is welcome, the focus on COVID-19 also threatens to obscure the fact that many researchers have long had to deal with fundamental uncertainty in their fieldwork. In fact, for those of us conducting research in settings of conflict or political repression, many of the challenges highlighted by the pandemic are not entirely new (Bond, Lake, and Parkinson 2020). There is a growing field of literature on how researchers can deal with different forms of uncertainty and risk, much of which originated in the "before times" (Glasius et al. 2018; Grimm et al. 2020; Krause and Szekely 2020; Mac Ginty, Brett, and Vogel 2020; Sriram 2009). While this literature does not directly address the pandemic, it deals with the management of uncertainty and risk in fieldwork.

A narrow focus on the challenges brought about by COVID-19 would neglect the fact that many forms of fieldwork have always been uncertain and risky. It would thus squander an important opportunity to recenter these issues. What is more, while countries in the Global North might be emerging from the crisis, global inequalities in vaccine access mean that the same is not true for large

parts of the Global South. Instead of discussing a "return to normal," we should therefore think seriously about how we can improve the way in which we deal with uncertainty and risk. This also includes taking a hard look at the stop-gap measures adopted in response to the pandemic, some of which might exacerbate risks in challenging settings.

While uncertainty in fieldwork is by no means only an effect of the pandemic, the COVID-19 crisis has demonstrated that our institutional risk assessment processes are lacking. Such mechanisms vary widely from country to country and institution to institution. Yet, there are several features which render these tools less than ideal.

First, risk assessment protocols are usually part of fieldwork authorization procedures of some form. Risk assessment might be one element of a larger ethical review framework conducted by IRBs, for example, or they might be stand-alone procedures implemented in the project design phase, or as part of travel authorization requirements. Frequently, however, such risk assessments set perverse incentives. Researchers depend on the ability to conduct research. This is especially true for early career researchers and those employed on precarious contracts. Risk assessments might thus appear as an additional administrative hurdle to be cleared in order to obtain coveted funding and researchers might be tempted to downplay risks. In other words, researchers are caught between two types of risk: the risk of being denied authorization with all its potential career consequences, and the risk associated with a specific research project. This situation, to put it mildly, is not conducive to a level-headed discussion about risks.

Second, formal risk assessment approaches are often based on crude distinctions between "dangerous" and "safe" fieldwork destinations and practices. A frequent procedure, for example, is to resort to travel advisories or risk maps provided by official institutions such as ministries of foreign affairs. Such risks maps are inappropriate tools for at least two reasons. First, they often do not sufficiently reflect sub-national variation in risk which can be of crucial importance, particularly to scholars conducting research in conflict zones. Second, risk maps are intended for very different purposes and cannot take the risk profiles associated with specific methodologies or research projects into account. They might therefore under- or overestimate the actual level of risk and are a far cry from the "nimble, consultative, and researcher-friendly" procedures advocated among fieldwork-active researchers themselves). Consequently, researchers are deprived of opportunities to learn from the experiences of others and are instead burdened with a largely bureaucratic exercise. As the pandemic has demonstrated, such procedures rarely produce contingency plans which are actually helpful in acute crisis situations.

Frustration with such bureaucratized procedures and with the fact that there was little by way of practical guidance on how to deal with risk in field research was one of the drivers behind the SAFEResearch initiative to which most of the authors in this symposium contributed. In a handbook emerging from this project (Grimm et al. 2020), we collected advice and best practices based on a collaborative process involving a large group of researchers, journalists, human rights defenders, and data activists. In the next section, I outline the risk assessment procedure we advocate there.

## Filling in a Form is Not Risk Assessment

Risk assessment is fundamental for both ethical and practical reasons. Ethically, the paramount do-no-harm principle requires a clear-eyed view of the potential risks for research participants as well as members of research teams. Practically, risk assessment is the basis for developing strategies to mitigate specific risks. Most researchers will care deeply about both of these issues, meaning that risk assessment will be important to them as well. The challenge is to organize risk assessment in a way that empowers research, rather than threatening to restrict it.

A research project's level of risk depends on the interaction between three sets of factors: the context in which research takes place, the nature of the project, and the characteristics of the research team itself. Looking at the problem this way should make it obvious why the simple red/yellow/green in risk maps is not an appropriate way of assessing risk.

I briefly illustrate these three elements:

*Risk assessment starts with context analysis.* For most social science researchers, understanding the political, historical, cultural, and social context of their prospective field sites starts long before field research. In fact, we frequently choose field sites *because of* specific features as part of larger research design and case selection strategies. But context analysis goes beyond general issues. It asks which specific actors are likely to oppose (or support) the research project, for what reasons, and resorting to what types of resources. Identifying such opposing and supporting actors can be difficult, and this stage is where researchers can profit from the experience of others, including other scholars, but also journalists, human rights workers, and other contacts in the field. Reaching out to such contacts in the design stage of a project is crucial as it allows researchers to develop an up-to-date view of the situation on the ground and to take fine-grained geographical variation in risk into

account. Especially where collaborators in the field are involved, it is sound advice to locate "decision making at the location of data, not the location of the principal investigator" (Scerri et al. 2020, 1572)—not only when it comes to assessing the public health situation on the ground. The concrete form of such processes will vary across research projects. While some researchers establish formal links to academic institutions in their field which give them access to local IRBs, this is not always possible, nor should it necessarily be seen as the gold standard. Requiring local IRB approval risks further bureaucratization, shifts administrative burdens from researchers' home institutions to those in the field, and might even be used as a mechanism to prevent politically sensitive research in some settings. Moreover, ethical review is still frequently seen as a one-off hurdle to be cleared. Instead, researchers should develop dynamic ways of feeding the perspectives of local collaborators into the risk assessment process. This empowers local collaborators, flattens hierarchies, and mitigates the risk of exploitative practices.

*No two research projects are the same.* While research as such might be seen with suspicion in some settings (Ahram and Goode 2016), assessments such as "location XYZ is dangerous" (or safe, for that matter) are usually not very helpful. The level of risk depends very much on what you are planning to do, including the sensitiveness of your research topic, the vulnerabilities of your research subjects, and the methods you are planning to use. Uncertain and risky field research does not only take place in the context of conflict or authoritarianism in the Global South, a message which has been driven home by the pandemic but was just as true before COVID-19.

*The strengths and weaknesses of research teams matters.* Fieldwork risk also depends on who you are and with whom you collaborate. Identity traits such as a researcher's nationality, ethnicity, gender, or sexual orientation might be characteristics to draw on in some settings but might increase vulnerability in others. The degree of local knowledge, including language skill and cultural competence, matters because it allows researchers to "read the room" and adapt to evolving situations. Working with trusted local partners as academic collaborators, research assistants, translators, or enumerators, is an important part of fieldwork for many of us. The input such local partners can provide is an invaluable part of risk assessment. Such relationships also generate their own risks, however. First and foremost, fieldwork safety refers to the safety of all research participants, importantly including local collaborators. In that sense, we can only mitigate risks to the extent that we mitigate them for the most vulnerable member of the research team. At the same time, especially in politically fraught

contexts, some collaborators might also pursue their own agendas and might thus bias our data collection process or even compromise the confidentiality of data or the anonymity of interlocutors.

Assessing risks in these three areas and the way they interact requires significant specialist knowledge. It requires an understanding of the field site and its political dynamics, of the methods and approaches to be used for data collection, and strong networks of trusted partners in the field. Not all of us can always rely on these crucial resources. Graduate students might prepare for their first fieldwork period and more experienced scholars might branch out to new field sites. This is when the ability to rely on the advice and experiences of others becomes important. In my experience, colleagues, NGO workers, and journalists on the ground are usually generous with their time and insights. Our institutional risk assessment procedures, however, do not encourage such networked approaches. In fact, risk assessment processes often take place under time pressure and research institutions do not usually provide the resources in terms of time, money, and training which are required for useful risk assessments.

## "Remote" and "Safe" Are Not Synonyms

Seeing their fieldwork disrupted during the pandemic, many of us were forced to innovate. Qualitative field research moved online, adding to the hours we all spent on Zoom during lockdowns (Howlett 2022). Messaging apps took on increasing importance as ways of staying in touch with our field. As political activity moved online as well, platforms such as WhatsApp, Twitter, or Telegram *became* the field for some researchers (Christia and Lawson 2020). While some elements of how we conduct fieldwork might revert "back to normal" as parts of the world emerge from the pandemic, others are here to stay.

Unfortunately, the stop-gap measures employed during the pandemic have not always accounted for the risks they involved. Rather, many forms of remote fieldwork involved an outsourcing of risk. I want to illustrate this with two examples. First, the pandemic has highlighted research practices which were considered problematic before. Most fieldwork projects rely on local collaborators and such forms of cooperation are not necessarily problematic. The potential for exploitative practices in such relationships has become particularly evident during the coronavirus crisis, however (Nyenyezi Bisoka 2020; Rudling 2021). Indeed, outsourcing research activities became the only feasible way forward for many research projects as international travel became impossible. The problem with this solution to the crisis is that it outsources risk along with research tasks. Given the fact that some local collaborators are

financially dependent on the research project, we cannot assume that their decisions are based on a disinterested risk assessment (Nyenyezi Bisoka 2020). In fact, this problem is not new, but is part of a larger problem in which narrowly understood risk assessment regimes shift risks away from principal investigators and onto local collaborators (Piccolino and Franklin 2019).

A second type of risk shifting involves moving research online. During lockdown, I used some of these tools myself and have conducted interviews and more informal conversations on Zoom, WhatsApp, and Telegram. There is nothing inherently wrong with having conversations online, yet as Jannis Grimm explains in more detail in this symposium, we must also recognize that we are shifting responsibility for risk mitigation to our conversation partners. In addition, few researchers fully understand the extent to which data exchanged on such platforms is protected against surveillance, an aspect that matters when conducting research in politically repressive settings. We might even be tempted to use platforms we know are not particularly safe (think Facebook Messenger, for example) because this is where our interlocutors are and asking them to move to a different platform might simply be asking too much.

There are no easy solutions to these problems. As we move toward a new normal in social science field research, however, we should continue to reflect on how we can make the process as safe as possible for all parties involved. A good place to start this process is to remove disincentives for clear-eyed risk assessment. Institutional risk assessment and clearance procedures are hard to change. If anything, their level of bureaucratization is likely to increase, particularly in Europe where such processes increasingly reflect principles adopted on the level of the European Union.[1] Proactively addressing issues of risk and safety in the design stage of a project can help researchers navigate these processes. It gives researchers agency in identifying risks and outlining mitigation procedures, while at the same time allowing for a thorough risk assessment process aimed at enabling safer research. Peer networks play an important role in this approach, and we should encourage further collaborative efforts on this issue. While this will not fully overcome the tension between the liability-driven risk aversity of institutions and researchers' own assessment of which risks might be worth taking, it at least allows researchers to use the process to their advantage. Admittedly, this is just a start; but we need to start somewhere.

## References

Ahram, Ariel I., and J. Paul Goode. 2016. "Researching Authoritarianism in the Discipline of Democracy." *Social Science Quarterly* 97, no. 4 (December): 834–49. https://doi.org/10.1111/ssqu.12340.

Baczko, Adam, and Gilles Dorronsoro. 2020. "The Ethical, Epistemological, and Conceptual Need to Resume Fieldwork." *Social Science Research Council: Items*, November 19, 2020. https://items.ssrc.org/covid-19-and-the-social-sciences/social-research-and-insecurity/the-ethical-epistemological-and-conceptual-need-to-resume-fieldwork/.

Bond, Kanisha D., Milli Lake, and Sarah E. Parkinson. 2020. "Lessons from Conflict Studies on Research during the Coronavirus Pandemic." *Items*, July 2, 2020. https://items.ssrc.org/covid-19-and-the-social-sciences/social-research-and-insecurity/lessons-from-conflict-studies-on-research-during-the-coronavirus-pandemic/.

Christia, Fotini, and Chappell Lawson. 2020. "Training the Covid-19 Cohort: Adapting and Preserving Social Science Research." *Items*, July 30, 2020. https://items.ssrc.org/covid-19-and-the-social-sciences/social-research-and-insecurity/training-the-covid-19-cohort-adapting-and-preserving-social-science-research/.

Fikrig, Kara. 2020. "I Evacuated My Fieldwork Site Because of COVID-19. What Now?" *Science*, May 7, 2020. https://doi.org/10.1126/science.caredit.abc6668.

Glasius, Marlies, Meta de Lange, Jos Bartman, Emanuela Dalmasso, Aofei Lv, Adele Del Sordi, Marcus Michaelsen, and Kris Ruijgrok. 2018. *Research, Ethics and Risk in the Authoritarian Field*. Cham: Springer International Publishing.

Grimm, Jannis, Kevin Koehler, Ellen 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 Publishing.

Howlett, Marnie. 2022. "Looking at the "Field" through a Zoom Lens: Methodological Reflections on Conducting Online Research during a Global Pandemic." *Qualitative Research* 22, no. 3 (June): 387-402. https://doi.org/10.1177/1468794120985691.

Krause, Peter, and Ora Szekely. eds. 2020. *Stories from the Field: A Guide to Navigating Fieldwork in Political Science*. New York: Columbia University Press.

Krause, Peter, Ora Szekely, Mia Bloom, Fotini Christia, Sarah Zukerman Daly, Chappell Lawson, Zoe Marks, et al. 2021. "COVID-19 and Fieldwork: Challenges and Solutions." *PS: Political Science & Politics* 54, no. 2 (April): 264–69. https://doi.org/10.1017/S1049096520001754.

1 See, for example, the Global Code of Conduct for Research in Resource-Poor Settings (Schroeder et al., n.d.) adopted by the European Commission in 2018.

Lake, Milli, and Sarah E. Parkinson. 2017. "The Ethics of Fieldwork Preparedness." *Political Violence @ a Glance*, June 5, 2017. https://politicalviolenceataglance.org/2017/06/05/the-ethics-of-fieldwork-preparedness/.

Mac Ginty, Roger, Roderick Leslie Brett, and Birte Vogel, eds. 2020. *The Companion to Peace and Conflict Fieldwork*. Cham: Palgrave Macmillan.

MacLean, Lauren M., Nabila Rahman, Robin Turner, and Jack Corbett. 2021. "Disrupted Fieldwork: Navigating Innovation, Redesign, and Ethics during an Ongoing Pandemic." *Qualitative & Multi-Method Research* 18, no. 2 (Fall): 1–8. https://doi.org/10.5281/ZENODO.4046546.

Nyenyezi Bisoka, Aymar. 2020. "Disturbing the Aesthetics of Power: Why Covid-19 Is Not an "Event" for Fieldwork-Based Social Scientists." *Social Science Research Council: Items*, May 28, 2020. https://items.ssrc.org/covid-19-and-the-social-sciences/social-research-and-insecurity/disturbing-the-aesthetics-of-power-why-covid-19-is-not-an-event-for-fieldwork-based-social-scientists/.

Piccolino, Giulia, and Sabine Franklin. 2019. "The Unintended Consequences of Risk Assessment Regimes: How Risk Adversity at European Universities Is Affecting African Studies." *Africa Spectrum* 54, no. 3 (December): 268–81. https://doi.org/10.1177/0002039719898904.

Rudling, Adriana. 2021. "Now Is the Time to Reassess Fieldwork-Based Research." *Nature Human Behaviour* 5 (8): 967. https://doi.org/10.1038/s41562-021-01157-x.

Scerri, Eleanor M. L., Denise Kühnert, James Blinkhorn, Huw S. Groucutt, Patrick Roberts, Kathleen Nicoll, Andrea Zerboni, et al. 2020. "Field-Based Sciences Must Transform in Response to COVID-19." *Nature Ecology & Evolution* 4 (12): 1571–74. https://doi.org/10.1038/s41559-020-01317-8.

Schroeder, Doris et al. n.d. "Global Code of Conduct for Research in Resource-Poor Settings." https://www.globalcodeofconduct.org/wp-content/uploads/2018/05/Global-Code-of-Conduct-Brochure.pdf

Sriram, Chandra Lekha, ed. 2009. *Surviving Field Research: Working in Violent and Difficult Situations*. London: Routledge.

Wood, Elisabeth Jean, Rogers Douglas, K. Sivaramakrishnan, and Rene Almeling. 2020. "Resuming Field Research in Pandemic Times." *Social Science Research Council: Items*, May 21, 2020. https://items.ssrc.org/covid-19-and-the-social-sciences/social-research-and-insecurity/resuming-field-research-in-pandemic-times/.

# The Mixed Blessing of Digital Fieldwork: Digital Security and Ethical Dilemmas of Remote Research during and after the Pandemic

Jannis Julien Grimm
*Freie Universität Berlin*

COVID-19 has markedly impacted the ways we collect research data through field research. As previously discussed in QMMR (MacLean et al. 2021) and elsewhere (e.g., GPPi 2021; ARC Bibliography 2021; SSRC 2020), the pandemic interrupted data collection and knowledge production routines. By restricting travel and free movement, thus impeding face-to-face exchanges, the pandemic and subsequent containment measures affected social scientists and their workflows, in particular those who previously relied on field-based methods. After all, interviews, ethnographic fieldwork, focus groups, and participant observation usually imply the physical co-presence of researchers and their participants, and often build on relations of trust that are established through repeated interpersonal contact. But quarantines, travel restrictions, lockdowns, social distancing, and even masks have made organizing personal encounters and maintaining and preserving dependable relations of trust with research participants harder—let alone establishing contact with and meeting new interlocutors.

At the same time, the pandemic has catalyzed the spread of old and the development of new online methodologies. The manifold ways in which COVID-19 disrupted qualitative research are outmatched only by the plethora of technical tools for "digital fieldwork" (see Digital Fieldwork 2021) adopted by researchers to compensate for lacking field access. It is safe to say that many of these new practices of doing fieldwork remotely are here to stay, even when field trips to large parts of the world become viable and ethically justifiable again. This is especially true for those technologies that

help translate traditional methods, such as interviewing or focus groups, to a virtual space, while at the same time being cheaper and less time intensive. Yet unlike the pandemic's toll on researchers and their projects and the impact of switching methods of data collection on the validity of the collected data, the corollary of an entire discipline "going digital" virtually overnight has been scarcely addressed in debates on field research during COVID-19.

## Digital Fieldwork and Remote Research: An Ambiguous Plan B

Commercial video chat services, social platforms and browser-based applications such as Zoom, Telegram, and Clubhouse became popular as tools for academic inquiry because they promised a quick way out of the bind faced by many researchers during the pandemic. Above all, the early days of COVID-19 were marked by a high degree of uncertainty about the viability, safety, or ethical permissibility of continuing research projects. In the context of a fast-spreading disease, many found that virtual research platforms held a comparative advantage: they relieved researchers of the duty to make a choice between going through with or aborting a planned project, or to draw the line, for instance, at a certain incidence rate. Given the potential health risks for scholars and their interlocutors posed by physical encounters in a situation of scarce information, individual exhaustion and uncertainty about the future, as well as emotional stress and anxiety, conducting interviews and ethnographies online (and thus in less obtrusive ways) even seemed to be the moral choice. Accordingly, many researchers moved their entire projects to virtual spaces as a mechanism to cope with the uncertainty of a rapidly changing pandemic situation. Often, they were encouraged to do so by advisors or supervisors who also found themselves unable to provide sound advice.

To master this shift from offline to online, scholars drew from a plethora of handbooks on virtual research techniques developed in the days before COVID-19 (Boellstorff et al. 2012; Braun, Clarke, and Gray 2017; Fielding, Lee, and Blank 2017; Hesse-Biber 2011; Kozinets 2010; Markham & Baym 2009; Abidin and de Seta 2020). This literature notwithstanding, moving online meant venturing into unfamiliar territory, fraught with risks and ethical dilemmas different to those which most researchers were already acquainted, including the intricacies of online data protection, the question of which archives may be legitimately mined as sources of primary data, the challenges of omnipresent surveillance for confidentiality, and epistemic questions about the power dynamics behind the knowledge produced in and from digital spheres (see Aldridge, Medina, and Ralphs 2010; van Baalen 2018; Grimm et al. 2020; Tanczer et al. 2020; Tanczer, McConville, and Maynard 2016; Rodham & Gavin 2006).

## Leveling the Playing Field?

The trend that social scientists "in the digital age" (van Baalen 2018, 2) were also increasingly reliant a digital-data infrastructure of which they often had only a rudimentary understanding was visible even before the COVID-19 outbreak (see Tanczer et al. 2020, 11). The latter only visualized this dependency in more obvious ways. Younger and more tech-savvy researchers especially compensated for a lack of field access by relying on virtual interviews, digital sources, or on local research assistants with whom, again, they communicated via various apps—some of which were launched in the wake of the pandemic.

The pandemic also highlighted how doing fieldwork had been a luxury often enjoyed by those with access to funding, support structures and training, and the right passport. At first, the pandemic seemed to level the playing field to a certain degree. With online surveys and focus groups, participant observations in Clubhouse or video chat sessions, discourse analyses of Facebook groups, the pandemic popularized a set of internet-based methodologies that offered time-sensitive access to research populations and that was equally available to well-funded researchers and those with less access to resources. In addition, the majority of the research community were newcomers when it came to these virtual research practices. This gave less privileged researchers the chance to catch up with the frontrunners. Within certain limitations, junior scholars also gained a certain advantage, because they often found it easier to adapt to and maneuver the new virtual research environment and the modes of data collection it entailed.

But the substitution of established data collection routines by new and often untested practices also catalyzed several worrying trends. In parallel to its partial equalization of research access, the methodological *tabula rasa*, above all, challenged central pillars of safe and ethical research conduct, including the principles of "informed consent" and "do no harm" in the relation between researchers and their research partners.

## Outsourcing Data Collection = Outsourcing Risk

Many researchers decided to opt not for a complete shift to digital methodologies but, instead, to rely on local research assistants (RAs) to compensate for the lack of field access. This option was certainly not available to everyone, as the employment of RAs is costly and not funded by every department. Still, the pandemic clearly

reinforced the preexisting trend towards the increased outsourcing of data collection to local researchers. To be clear, the employment of local RAs is not per se an exploitative practice. It often simply reflects the exigencies and constraints at different points in someone's life or career: senior scholars have less time to go on long field trips, and for emerging junior researchers, serving as RAs can be an important stepping-stone in their careers. However, these practices always risk replicating an unequal distribution of burden and merit between both sides involved in the transaction (see Eriksson Baaz & Utas 2019). And the increasing resort to digital technologies tends to conceal this imbalance further. For centuries, the social sciences have maintained a questionable record of rendering local researchers and their work invisible. As Aymar Nyenyezi Bisoka (2020) has noted: "When the time comes for 'difficult' fieldwork in Africa, research assistants become body-instruments, an extension of the bodies of Global North researchers."

Today, Global North researchers are no longer carried on the backs of locals through swamplands, but they do depend more than ever on local experts to maneuver the difficult terrain created by the pandemic. This outsourcing of data collection and analysis also entails the outsourcing of the potential risks incurred during the research process. When projects rely on a complex layer of digital infrastructure as a mediating mechanism for locally collected data, it is not only the local data collectors who are rendered invisible, but also the difficult terrain they face. Not only that, the digital infrastructure itself becomes a potential source of hazard; when projects depend on technology for communication, storage, and joint analysis of primary data, this also means that the burden of coming up with safe internet connections, secure communication technologies, and ways of safely storing data is increasingly placed on local RAs. This is particularly problematic in heavily surveilled field sites where security apparatuses keep a close eye on researchers' communications. Such "hostile environments" call for discretion and for the creation of less, not more, data files, phone records, online paper trails, and other communication at risk of interception (see Mwambari, Purdeková, and Nyenyezi Bisoka 2021, 3), especially if they are produced by people who don't have the option to leave the field site, or for whom "the field" starts right at their doorsteps.

This aspect is often insufficiently considered by project leaders, but also by IRBs and editorial boards. While ethical review boards are usually clear about the measures required of researchers to protect their informants during the collection of primary data, they often don't interrogate how the outsourcing practices may engrain an asymmetric distribution of risks and

merits into partnerships between researchers and their RAs, and by extension, between academics in the Global North and Global South. The cessation of physical travel to a field site may allow the former to avoid health risks and help contain a global pandemic, but if the halted onsite fieldwork is simply outsourced, it can expose the latter to even more to dangerous conditions.

## Spatial Separation and Affective Detachment

In addition to increasing asymmetries, digital field work more generally adds a barrier to affective solidarity with the subjects of our inquiries. Remote research simply makes real world problems more remote.

First, it affects the sensitivity of academic researchers towards the everyday needs of their interlocutors. As Kanisha Bond, Milli Lake, and Sarah Parkinson (2020) noted, "a rush to conduct face-to-face surveys with distressed populations; to monetarily incentivize interviews in victimized communities; or to otherwise collect political data from individuals without critical evaluations of the social and scientific urgency of such work greatly risks elevating researcher priorities over research participants' current needs." The employment of remote research techniques that are less conditioned by institutional approval procedures is bound to facilitate these practices.

Second, it affects researchers' awareness of the threats that research participants may be exposed to, an interferes with their ability to "feel" a place. Especially for those that are new to the study of a certain country or context, remote methodologies make it incredibly hard to estimate the risks of their interventions. Supervisors and senior researchers are also less able to provide advice in this situation as they are often not sufficiently acquainted with the technologies used by their students. Consequently, in a fully digitalized, remote research environment, it is much harder for them to assist and to fulfil their duty of care than during classical field research. Mwambari et al. (2021) made a similar argument highlighting how remote research impedes researchers' immersion into the field as well as the trust-building and context awareness that are so essential for planning and conducting safe and sensitive research.

Third, it also concerns researchers' sensitivity to the threat of surveillance. One major problem with the surveillance of digital communications technologies is that it is by definition hard to detect. This problem is exacerbated by many researchers' lack of technical savvy. Every day we make lots of passive, habitual or accidental, decisions that we are not aware of when we use technologies, for instance, by clicking "agree" in some popup window and thereby passively confirming

a user agreement and privacy policy that we neither read nor understand. Not all of these habitual decisions may expose our work to surveillance. But some do, and the predicament is that most of us are in no position to discern which. Consequently, it is hard to calculate the risk from surveillance that researchers and their interlocutors are facing when engaging with each other exclusively through technological media. What is more, the privacy infringements that may result from miscalculation are likely to remain unknown and untraceable (Aldridge, Medina, and Ralphs 2010, 3), thus impeding collective or institutional learning from experiences.

When research institutions approve the use of moderated Clubhouse sessions for focus group discussions, or a specific video software for interviewing, the motives behind this are often laudable. It stems from their realization of a duty of care towards their staff and students, and it often is an attempt to enable safer alternatives than travelling and group meetings during a pandemic. But these calculations often fail to take into account how technologies can also be instrumentalized for intimidation, repression or surveillance—such as when records of video-chats are leaked online to expose participants or when members of Clubhouse sessions are intimidated on-air (Iskandarani 2021). Not only will these abuses of technology remain unknown to those in charge of approval procedures at the institutional level, these individuals are probably not equipped to provide practical advice on handling such incidents.

It is imperative that we start reflecting about ways to reform established institutional practices so as to make them best serve an increasingly datafied profession. This includes pre-field work courses, departmental workflows, and staff training. Where the necessary workflows and support structures at academic institutions are lacking, we should try to develop these in-house capacities and tailor them to the risks that come with increasing technological dependence, for instance, by installing departmental focal points for communication and data security, or by integrating mandatory digital risk assessments into the field work approval process.[2]

## Conclusion

What is worrying is not the shift towards more technology-assisted analysis—an inevitable trend that was there before the outbreak of COVID-19 is bound to continue. Rather, what is worrying is that the increasing dependence on little-known technologies has not yet led to a greater emphasis on digital literacy education within the profession. Even after one and a half years of webinars, online lectures, virtual focus groups and the like, we still lack compelling answers to the question of how we should deal with surveillance and untransparent technologies. When it comes to the allocation of resources, personnel, and training, we mostly treat digital security as a side-aspect, rather than an integral part of project planning. We still prefer to ignore well-known privacy concerns about certain platforms and tools, when we should instead aspire to understand them better. Even less time has been spent on discussing the practical implications of norms such as do no harm and informed consent for the digital research "on steroids" that we have witnessed since the start of the pandemic.

Before the outbreak of COVID-19, projects like SAFEResearch (Grimm et al. 2020) aimed to come up with practical guidance for researchers on how to take informed and ethical decisions when (re)designing the field work stages of their projects. Crucially, this included the issue of digital security and data protection (89-127). The project aimed to move the discipline from a passive observation of its increasing dependence on little-understood software to a more active decision-making process on the use of technology. Building on this and similar guidance, we should identify and share good practices for moving research online and dealing with ethical ramifications during and after the pandemic. This doesn't mean we all need to become digital security experts. But we should aim to become more "digitally literate," to know whom and what to ask, and move from passive towards more active decisions on research technologies—that is to say, conscious decisions which are planned and based on a reliable degree of information.[3] Otherwise, we risk undermining the very ethical frameworks we cherish. After all, how much informed consent is possible, if we neither have the ability to grasp nor to comprehensively inform our research partners about the potential risks of their participation?

Unfortunately, this may entail accepting that to some of the most convenient tools that we have grown used over the last years, are potentially the least adequate for research purposes from a safety and ethics perspective. But it also means that we cannot simply try out every new app or software that becomes available for our research in the hope that we might come across a tool that is both safe and easy to use. Pandemic or not, most of us are doing research with real people. Their lives are not a testing ground for new methodologies and software.

---

2 Kevin Koehler has written in more detail about risk assessments as a means to empower, rather than constrain researchers this symposium. For a series of templates for digital risk assessments see Grimm et al. (2020, 94–95, 106, 126-127).

3 Sebastian Van Baalen, METAReSPS Roundtable on Digital Security and Data Protection (University of Bologna, Bologna, Italy, May 12, 2021).

Finally, we should stop treating the pandemic as a logistical challenge to be overcome through sophisticated toolkits and the outsourcing of risk. Social science field research has always been defined by unequal burden-sharing between local and foreign knowledge producers. Once we acknowledge how the shift to remote methodologies feeds into this asymmetric relationship, we can start having a more productive conversation on the parameters and incentive structures needed to steer the evolution of research practices in ways that facilitate the safe and ethical conduct of social inquiry.

## References

Abidin, Crystal, and Gabriele de Seta, eds. 2020. "Doing Digital Ethnography: Private Messages from the Field," special issue, *Journal of Digital Social Research*. 2, no. 1 (February). https://jdsr.se/ojs/index.php/jdsr/issue/view/3

Aldridge, Judith, Juanjo Medina, and Robert Ralphs. 2010. "The Problem of Proliferation: Guidelines for Improving the Security of Qualitative Data in a Digital Age." *Research Ethics 6*, no. 1 (March): 3–9. https://doi.org/10.1177/174701611000600102.

ARC Bibliography. 2021. "Research During the Coronavirus Covid-19 Pandemic." *Advancing Research on Conflict,* https://advancingconflictresearch.com/researchincrisis

Boellstorff, Tom, Bonnie Nardi, Celia Pearce, and T.L. Taylor. 2012. *Ethnography and Virtual Worlds: A Handbook of Method*. Princeton: Princeton University Press.

Bond, Kanisha. D., Milli Lake, and Sarah E. Parkinson. 2020. "Lessons from Conflict Studies on Research during the Coronavirus Pandemic." *SSRC Items*, July 2, 2020. https://items.ssrc.org/covid-19-and-the-social-sciences/social-research-and-insecurity/lessons-from-conflict-studies-on-research-during-the-coronavirus-pandemic/

Braun, Virginia, Victoria Clarke, and Debra Gray, eds. 2017. *Collecting Qualitative Data: A Practical Guide to Textual, Media and Virtual Techniques*. Cambridge: Cambridge University Press.

Eriksson Baaz, Maria, and Mats Utas. 2019. "Exploring the Backstage: Methodological and Ethical Issues Surrounding the Role of Research Brokers in Insecure Zones." *Civil Wars* 21 (2): 157–78. https://doi.org/10.1080/13698249.2019.1656357.

Fielding, Nigel, Raymond M. Lee, and Grant Blank. eds. 2017. *The SAGE Handbook of Online Research Methods*. Thousand Oaks: SAGE Publishing.

GPPi. 2021. "COVID-19 and Academic Freedom: What's the Connection?" *GPPi (Global Public Policy Institute),* September 2021. https://covidacadfreedom.gppi.net.

Grimm, Jannis, Kevin Koehler, Ellen 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 Publishing.

Hesse-Biber, Sharlene N. ed. 2011. *The Handbook of Emergent Technologies in Social Research*. Oxford: Oxford University Press.

Iskandarani, Aya. 2021. "Lebanon: Hezbollah critics report threats on Clubhouse app." The National, April 21, 2021. https://www.thenationalnews.com/mena/lebanon/lebanon-hezbollah-critics-report-threats-on-clubhouse-app-1.1200400.

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

Kozinets, Robert. V. 2010. *Netnography: Doing Ethnographic Research Online*. Thousand Oaks: SAGE Publications.

MacLean, Lauren, Nabila Rahman, Robin Turner, and Jack Corbett. 2021. "Disrupted Fieldwork: Navigating Innovation, Redesign, and Ethics during an Ongoing Pandemic." *Qualitative & Multi-Method Research* 18, no. 2 (Fall): 1–8. https://doi.org/10.5281/zenodo.4046546.

Markham, Anette N., and Nancy K. Baym. 2009. *Internet Inquiry: Conversations About Method*. Thousand Oaks: SAGE Publications.

Mwambari, David, Andrea Purdeková, and Aymar Nyenyezi Bisoka. 2021. "Covid-19 and Research in Conflict-Affected Contexts: Distanced Methods and the Digitalisation of Suffering." *Qualitative Research*, online first. https://doi.org/10.1177/1468794121999014

Nyenyezi Bisoka, Aymar. 2020. "Disturbing the Aesthetics of Power: Why Covid-19 Is Not an "Event" for Fieldwork-based Social Scientists." *Social Science Research Council: Items*, May 28, 2020. https://items.ssrc.org/covid-19-and-the-social-sciences/social-research-and-insecurity/disturbing-the-aesthetics-of-power-why-covid-19-is-not-an-event-for-fieldwork-based-social-scientists/

Rodham, Karen, and Jeff Gavin. 2006. "The Ethics of Using the Internet to Collect Qualitative Research Data." *Research Ethics* 2, no. 3 (September): 92–7. https://doi.org/10.1177/174701610600200303

SSRC. 2020. "Covid-19 and the Social Sciences." *Social Science Research Council: Items*, April 25, 2020. https://items.ssrc.org/category/covid-19-and-the-social-sciences/

Tanczer, Leonie Maria, Ronald J. Deibert, Didier Bigo, M. I. Franklin, Lucas Melgaço, David Lyon, Beky Kazansky, and Stefania Milan. 2020. "Online Surveillance, Censorship, and Encryption in Academia." *International Studies Perspectives* 21, no. 1 (February): 1–36. https://doi.org/10.1093/isp/ekz016

Tanczer, Leonie Maria, Ryan McConville, and Peter Maynard. 2016. "Censorship and Surveillance in the Digital Age: The Technological Challenges for Academics." *Journal of Global Security Studies* 1, no. 4 (November): 346–55. https://doi.org/10.1093/jogss/ogw016.

van Baalen, Sebastian. 2018. "'Google wants to know your location': The ethical challenges of fieldwork in the digital age." *Research Ethics* 14, no. 4 (October): 1–17. https://doi.org/10.1177/1747016117750312

# Reflexive Advising: Engaged Mentorship for Safe and Ethical Research Practice

**Sarah E. Parkinson**
*Johns Hopkins University*

**Dina Zayed**
*University of Sussex*

Political science arrived comparatively late to conversations regarding fieldwork safety. Professional fields such as journalism and humanitarian aid began providing practical training—including first aid and risk assessment strategies—to employees deployed to violence-affected, repressive, and unstable contexts starting in the 1990s (Lake and Parkinson 2017). Like their colleagues in other fields, academics often travel to remote places, examine contentious topics, and rely heavily on local buy-in for access and safety. Even as research in such sites has increased, many scholars report feeling practically unprepared for their fieldwork (Cronin-Furman and Schwartz 2020).

The COVID-19 pandemic, along with other dynamics, has shifted the possibilities for academic fieldwork, rendering "the field" more uncertain for many researchers. Vaccine inequality makes travel to places with high caseloads or low vaccine access ethically fraught. Moreover, the recent targeting of foreign academics such as Matthew Hedges (Siddique 2021), as well as a long-standing trend in many countries of intimidating local scholars (Human Rights Watch 2018; 2021; Kaczmarska and Dubrovsky 2020) demonstrate the need for institutionalized safety practices and education. Yet uncertainty over events and conditions and concerns regarding how to negotiate the late-/post-pandemic context force scholars to ask: How can we develop "best practices" in a realm where there are often no right answers? How do we institutionalize robust research behavior when even well-trained and resourced scholars find themselves in potentially unsafe situations?

While pedagogical literature is beginning to surface on the ethics surrounding advising in graduate training (see, e.g., Eck and Cohen 2020), emergent work around the conduct of ethical research rarely touches upon a cornerstone of academic socialization embodied in the relationship between adviser and advisee. This essay thus argues that what we term "reflexive advising"—where mentors and mentees collectively acknowledge and evaluate how their positionalities may shape their research experiences—contributes a useful additional framework to more individualized, existing approaches to practicing researcher safety. We argue that this relationship is a crucial grounding point for ethical conversations, and one which must center ethical thought and dialogical learning for research design and practice.

Drawing on the authors' experiences working with the Advancing Research on Conflict (ARC) Consortium (Parkinson) and with the Research Ethics in the Middle East and North Africa (REMENA) project (Parkinson and Zayed),[4] the remainder of this essay broadly outlines researcher safety concerns, then presents an outline for reflexive advising, which responds to emergent calls for a shift to re-balance the burden of safe research between early-career researchers and their mentors.

## Relational Risk in the Field

Significant attention has been placed on the ethical and methodological challenges of fieldwork in repressive, violence-affected, and fragile settings (see, e.g., Wood 2003, 2006; Fujii 2010; Ahram and Goode 2016; Campbell 2017; Glasius et al. 2018; Knott 2019; Ryzova 2017; Grimm et. al. 2020; Krause 2021;

---

4 The Advancing Research on Conflict (ARC) Consortium was founded in 2018 to foster methodologically robust, ethical, context-sensitive research on conflict and violence. The REMENA Project (REMENA 2020) is "dedicated to mobilizing an interdisciplinary network of academics, researchers and practitioners to assess the landscape of social science research conducted in the Arab world and develop guidelines for the conduct of responsible, ethical and constructive social inquiry."

Parkinson 2022). In some settings, political dynamics and misunderstanding of research activities may even position researchers as suspected "agents of conspiracy" (Sowers 2015; Sholkamy 2015; Holmes and Aziz 2019; Driscoll and Schuster 2017). This discussion has become even more robust in the contest of the COVID-19 pandemic, which has brought questions of risk, inequality, power, surveillance, and exploitation in both field research and higher education as a whole to the fore (see, for example, the Social Science Research Council's essay forum on COVID-19 and the Social Sciences). However, little work addresses how to responsibly advise students and how to evaluate in advance whether proposed field-based work appropriately weighs safety considerations.

In circumstances of hyper-nationalism and under backdrops where academics are expected to pursue clearances for virtually all their activities, conducting even seemingly unimposing research can be risky. A close reading of a range of recent cases of researchers who have come into uncomfortable contact with security agencies indicates that the circumstances through which they faced harassment or worse were often triggered by suspicions or professed suspicions from their research participants. In Egypt, the recent arrest of two postgraduate students from Alexandria University who were conducting a survey on the use of the city's public transportation network is indicative. After one passenger reported his concerns that the pair might be plotting a "terrorist attack" to a local police officer, Egyptian authorities detained the students for investigation. Of course, such mobilization of suspicion is enabled by the regime's self-serving and self-perpetuating infrastructure of policing, where citizens are encouraged to partake in reproducing the power of the security establishment (Abdelrahman 2017).

Especially for inexperienced researchers, knowledge of risks in the field may shape a heavy emphasis on methodological considerations around conducting fieldwork in a manner that both guarantees access and avoids security backlash. However, at the heart of these risks are ethical questions around how to design and conduct research in a manner that avoids harm and heeds the principle of beneficence. Researchers develop what Yanow & Schwartz-Shea (2018) characterize as a "situational persona" to negotiate access and maintain trust with research participants. They routinely occupy "positional spaces," whereby both researchers and their collaborators are "involved in a constant shifting of the multiple axes" upon which identities rest (Mullings 1999, 341; also see Narayan 1993). It is widely understood that methodological techniques reliant on networks of trust

also expose researchers, participants, and team members alike to surveillance, and potentially reprisal (Clark 2006; Tripp 2018). Accessing many fields sites relies upon introductions through trusted peers and institutional partners in an enactment of associational trust; the trust placed by interviewees and research interlocutors in their own contacts determines the trust that may be granted to interviewers themselves (Malejacq and Mukhopadhyay 2016).

How can advisers support advisees in considering ethical obligations and safety concerns that are wed to and reproduced by the dynamics of their field sites? In what ways can advisers support advisees in thinking about the dynamic evolution of their "situational personas" in evolving field conditions?

## Reflexive Advising as Engaged Risk Assessment

Advising is integral to the training and socialization of academic researchers. Productive advising constitutes a simultaneous act of supporting students and serving as a "trustworthy gatekeeper" as the "principal interface" between students, departments, and disciplines (Barnes and Austin 2009; Schlosser et al. 2011; Vacha-Haase et al. 2004).

Advising is rarely politically neutral; it requires pedagogical grounding that promotes what Puroway (2016, 4) describes as "critical reflection and action consistent with praxis." As Schlosser et al. (2011, 55) note, advisers are prompted to understand the influence of "culture, race, and ethnicity, as well as, one's level of acculturation, enculturation" in engaging with multicultural advising relationships. Advising is itself a site of ethical practice (Lee & Metcalfe 2017).

Yet, empirical research (Cronin-Furman and Schwartz 2020)[5], as well as discussions we've had with other members of ARC and REMENA, reveal a stark reality: many graduate students and junior scholars in North American and European academic institutions conduct research without a robust mentorship network and with little safety net beyond their own instinct. Despite the significant role advisers play in shaping graduate research projects, in practice, many advisors remain in the dark regarding the ethical and safety challenges their mentees face in the field.

How can advisors who do not have a similar research background guide students? How can advisors with comparable fieldwork backgrounds productively leverage their experiences across contexts for advising purposes? How can advisors best evaluate student competencies? When do advisors need to develop new competencies,

5 Cronin-Furman and Schwartz (2020) surveyed international relations and comparative politics students and faculty regarding their views on and training for international fieldwork.

encourage a mentee to broaden their advising team (e.g., by including someone with in-country experience where they are headed), or, quite simply, say "no"? We suggest a system whereby mentors and mentees actively and collectively consider a combination of researcher positionality and contextual factors in order to open discussions of field safety. As a short, initial example of such an evaluation, consider the potential difference in positionality between the following five students proposing research in Morocco:

1) An American student who entered a US-based PhD program directly from undergraduate who has no non-English language skills and no prior Middle East travel;

2) An American student who is fluent in Arabic, studied abroad in Jordan, and entered a US-based PhD program after working in the Middle East for four years;

3) A Moroccan national who entered a US-based PhD program following a decade of professional experience in international organizations in North Africa;

4) A Moroccan national who entered a US-based PhD program and plans to use their family's government connections to access a restricted area;

5) A dual US-Moroccan national who entered a US-based PhD Program following a decade of working for the US Department of State.

There exist objective and subjective differences between these researchers, especially when we consider how they might be treated in the field by their interlocutors, any research team members, the Moroccan state, and the US government. There are ways for each to conduct robust, ethical research in Morocco. Each of these students deserves engaged, thoughtful advising. Yet each will also enter the field with a very different positionality and distinct safety concerns; each requires tailored advising. The student with no prior experience might unknowingly stumble into politically sensitive entanglements, raising state suspicion of their motivations. A Moroccan national might have in-country family and friends who could be endangered if they were perceived as criticizing the monarchy. By contrast, a Moroccan with family connections in the government might be perceived as a threat to participants (Research Ethics in Kashmir 2021). A former US government employee, by relying on their past professional connections or professional mannerisms, might unintentionally act in ways that mimic those of their past position and compromise consent from interviewees.

The goal of reflexive advising is to consider potentially relevant aspects of researcher positionality in advance and to evaluate how they may influence their encounters in the field. While the exercise above highlights specific aspects of positionality related to researcher identity, reflexive advising may also include questions regarding, for example, sources of research funding or histories of extractive research at field sites. The advisor's own positionality in these interactions also matters. For example, an advisor with in-country experience might understand the project's risks in very different ways than one with regional experience or no background in the region. An adviser's own reputation—given, for example, social media posts or a public history of US defense consulting—can also shape the perception of a mentee in the field. It is here that regional and topic-oriented groups as well as professional associations—in our cases, organizations such as the Project on Middle East Political Science (POMEPS), the Middle East Studies Association (MESA), the Arab Council for the Social Sciences (ACSS), and the REMENA project—provide crucial mentorship and feedback opportunities for early-career researchers beyond their immediate advising network.

It is also essential to consider how an advisor's own biases about the context being studied or about the researcher may shape advising in problematic (and potentially discriminatory) ways. For example, both of this article's authors have been told that their projects were so dangerous as to be unworkable. In several cases, such statements have been based on flawed understandings of gender dynamics in the Middle East. It must be unequivocally stated: there exists a massive difference between reflexive discussion of researcher positionality and the racist/sexist stereotyping of researchers and the spaces that they study. Advisors are responsible for educating themselves and acknowledging their own biases throughout the advising process.

## Conclusion

Reflexive advising invites mentors and mentees to simultaneously assess their knowledge, capacities, and potential blind spots. It demands that they acknowledge the broader politics of research and situate the scholar within them. The idea is not to say that specific types of people should never do certain types of research or should not work in certain sites. Rather, the primary idea behind reflexive advising is to encourage scholars to acknowledge that researchers' positionalities inevitably shape their experiences in the field and to initiate relevant advising conversations. The exercise of considering researcher positionality at the beginning of a project reveals a deeper logic of honest communication and robust research design that derive from both common sense and established individual-level practices.

# References

Abdelrahman, Maha. 2017. "Policing Neoliberalism in Egypt: The Continuing Rise of the 'Securocratic' State." *Third World Quarterly* 38 (1): 185–202. https://doi.org/10.1080/01436597.2015.1133246.

Ahram, Ariel. I. and J. Paul Goode. 2016. "Researching Authoritarianism in the Discipline of Democracy." *Social Science Quarterly,* 97, no. 4 (December): 834-49. https://doi.org/10.1111/ssqu.12340.

Barnes, Benita J. and Ann E. Austin. 2009. "The Role of Doctoral Advisors: A Look at Advising from the Advisor's Perspective." *Innovative Higher Education* 33: 297–315. https://doi.org/10.1007/s10755-008-9084-x.

Campbell, Susanna P. 2017. "Ethics of Research in Conflict Environments." *Journal of Global Security Studies* 2, no. 1 (January): 89–101. https://doi.org/10.1093/jogss/ogw024.

Clark, Janine A. 2006. "Field Research Methods in the Middle East." *Political Science and Politics* 39, no. 3 (July): 417-24.

Cronin-Furman, Kate, and Schwartz, Stephanie. 2020. "Ill-Prepared: International Field Methods Training." Unpublished manuscript, London, UK.

Driscoll, Jesse, and Schuster, Caroline. 2018. "Spies like us." *Ethnography* 19, no. 3: 411–430. https://doi.org/10.1177/1466138117711717.

Eck, Krsitine, and Cohen, Dara Kay. 2021. "Time for a Change: The Ethics of Student-Led Human Subjects Research on Political Violence." *Third World Quarterly* 42 (4): 855-66. https://doi.org/10.1080/01436597.2020.1864215.

Fujii, Lee Ann. 2012. "Research Ethics 101: Dilemmas and Responsibilities." *PS: Political Science & Politics* 45, no. 4 (October): 717–23. https://doi.org/10.1017/S1049096512000819.

Glasius, Marlies, Meta de Lange, Jos Bartman, Emanuela Dalmasso, Aofei Lv, Adele Del Sordi, Marcus Michaelsen, and Kris Ruijgrok. 2018. *Research, Ethics and Risk in the Authoritarian Field.* Basingstoke: Palgrave Macmillan.

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 Ltd.

Holmes, Amy A. and Sahar Aziz. 2019. "Egypt's Lost Academic Freedom." *Sada Middle East Analysis.* January 24, 2019. https://carnegieendowment.org/sada/78210.

Human Rights Watch. 2018. "Turkey: Government Targeting Academics." *Human Rights Watch.* May 14, 2018. https://www.hrw.org/news/2018/05/14/turkey-government-targeting-academics.

Human Rights Watch. 2021. "'They Don't Understand the Fear We Have': How China's Long Reach of Repression Undermines Academic Freedom at Australia's Universities." *Human Rights Watch.* https://www.hrw.org/report/2021/06/30/they-dont-understand-fear-we-have/how-chinas-long-reach-repression-undermines.

Katarzyna, Katarzyna, and Dmitry Dubrovsky. 2020. "Russia's Limits on Critical Thinking Are Hitting Its Academic Performance." *Times Higher Education,* September 13, 2020. https://www.timeshighereducation.com/blog/russias-limits-critical-thinking-are-hitting-its-academic-performance.

Knott, Eleanor. 2019. "Beyond the Field: Ethics after Fieldwork in Politically Dynamic Contexts." *Perspectives on Politics* 17, no. 1: 140–53. https://doi.org/10.1017/S1537592718002116.

Krause, Jana. 2021. "The Ethics of Ethnographic Methods in Conflict Zones." *Journal of Peace Research,* 58, no. 3 (May): 329-341. https://doi.org/10.1177/0022343320971021.

Lake, Milli, and Sarah E. Parkinson. 2017. "The Ethics of Fieldwork Preparedness." *Political Violence @ a Glance.* June 5, 2017. http://politicalviolenceataglance.org/2017/06/05/the-ethics-of-fieldwork-preparedness/.

Lee, Yuko Igekami, and Amy Metcalfe. 2017. "Academic Advisors and Their Diverse Advisees: Towards More Ethical Global Universities." *Journal of International Students* 7 (4): 944-62. https://doi.org/10.32674/jis.v7i4.183.

Malejacq, Romain, and Dipali Mukhopadhyay. 2016. "The 'Tribal Politics' of Field Research: A Reflection on Power and Partiality in 21st-Century Warzones." *Perspectives on Politics* 14, no. 4 (December): 1011–28. https://doi.org/10.1017/S1537592716002899.

Mullings, Beverley. 1999. "Insider or Outsider, Both or Neither: Some Dilemmas of Interviewing in a Cross-Cultural Setting." *Geoforum* 30, no. 4 (November): 337–50. https://doi.org/10.1016/S0016-7185(99)00025-1.

Narayan, Kirin. 1993. "How Native Is a 'Native' Anthropologist?" *American Anthropologist* 95, no. 3: 671–86. https://doi.org/10.1525/aa.1993.95.3.02a00070.

Parkinson, Sarah E. 2022. "(Dis)Courtesy Bias: 'Methodological Cognates,' Data Validity, and Ethics in Violence-Adjacent Research." *Comparative Political Studies* 55, no. 3 (March): 420-50. https://doi.org/10.1177/00104140211024309.

Puroway, Andrew W. 2016. "Critical Advising: A Freirian-Inspired Approach." *NACADA Journal* 36 (2): 4-10.

REMENA Project. 2020. "Call for Participation: Workshops on Research Ethics in the Middle East and North Africa." Middle East Institute, Columbia University. https://www.mei.columbia.edu/news-fall-2020/remena-call-for-participation.

Research Ethics in Kashmir. 2021. "Statement on Anthropologist Saiba Varma." *Medium*. September 20, 2021. https://medium.com/@ResearchethicsinKashmir/statement-on-anthropologist-saiba-varma-f1bd0474d470.

Ryzova, Lucie. 2017. "New Asymmetries in the New Authoritarianism: Research in Egypt in the Age of Post-Revolution." *International Journal of Middle East Studies* 49, no. 3 (August): 511-514. https://doi.org/10.1017/S002074381700037X.

Schlosser, Lewis Z., Heather Z. Lyons, Regine M. Talleyrand, Bryan S. K. Kim, and W.Brad Johnson. 2011. "A Multiculturally Infused Model of Graduate Advising Relationships." *Journal of Career Development* 38, no. 1 (February): 44-61. https://doi.org/10.1177/0894845309359286.

Sholkamy, Hania. 2015. "Egypt's Love-Hate Relationship with Research." *Ahram Online*, June 26, 2015. http://english.ahram.org.eg/NewsContentP/4/133755/Opinion/Egypts-lovehate-relationship-with-research.aspx.

Siddique, Haroon. 2021. "British Academic Sues UAE Officials over Torture Claims While in Detention." *The Guardian*, May 6, 2021. https://www.theguardian.com/world/2021/may/06/british-academic-sues-uae-officials-for-alleged-torture-while-in-detention-matthew-hedges.

Sowers, Jeannie. 2015. "Activism and Political Economy in the New–Old Egypt." *International Journal of Middle East Studies* 47, no. 1 (February): 140–143. https://doi.org/10.1017/S0020743814001500.

Tripp, Aili Mari. 2018. "Transparency and Integrity in Conducting Field Research on Politics in Challenging Contexts." *Perspectives on Politics* 16, no. 3 (September): 728–38. https://doi.org/10.1017/S1537592718001056.

Vacha-Haase, Tammi, Donna S. Davenport, and Shoshana D. Kerewsky. 2004. "Problematic Students: Gatekeeping Practices of Academic Professional Psychology Programs." *Professional Psychology: Research and Practice* 35, no. 2 (April): 115-22. https://doi.org/10.1037/0735-7028.35.2.115.

Wood, Elisabeth Jean. 2006. "The Ethical Challenges of Field Research in Conflict Zones." *Qualitative Sociology* 29 (June): 373–86. https://doi.org/10.1007/s11133-006-9027-8.

——. 2003. *Insurgent Collective Action and Civil War in El Salvador.* Cambridge: Cambridge University Press.

Yanow, Dvora, and Peregrine Schwartz-Shea. 2018. "Framing "Deception" and "Covertness" in Research: Do Milgram, Humphreys, and Zimbardo Justify Regulating Social Science Research Ethics?" *FQS: Forum Qualitative Social Research* 19 (3): Art. 15. https://doi.org/10.17169/fqs-19.3.3102.

# Making Sense of Corruption Prosecutions in Peru: Metanarratives as a Tool for Interpreting Elite Interviews

Will Freeman
*Princeton University*

The politician belonged to a party that was a front for one of Peru's biggest criminal organizations—at least according to public prosecutors. That's why I was surprised when she agreed to an interview, and suggested we meet at a sunny, open-air café in an upscale neighborhood of the capital, Lima. I had come to Peru to study how states end impunity for grand corruption. Little about Peru's politics or history predicted a strong, independent judiciary. But since 2016, a small, underfunded team of public prosecutors had put dozens of politicians long considered above the law behind bars. Having already interviewed over a dozen prosecutors

and judges, I was eager to get through to the other side: Peru's once-powerful political class, which had failed spectacularly in its attempts to halt the investigations. Now was my chance.

Stuck in Lima's notorious bumper-to-bumper traffic, I wondered what to expect. The politician I was meeting hailed from Fujimorismo, a right-wing populist movement that emerged from the ashes of the competitive authoritarian regime of ex-President Alberto Fujimori (1990-2000). Despite the regime's record of human rights abuses, murder, and graft, the Fujimorista party, Popular Force, remained one of Peru's most powerful political parties until the corruption investigations started in 2016. Since then, prosecutors had arrested several senior Fujimorista politicians, including Keiko Fujimori, the daughter of the ex-dictator and the leader of Popular Force. Prosecutors were investigating the party itself for allegedly acting as a vehicle for money laundering and organized crime.

I had made a habit of scouring the web for past interviews with my sources, which tended to provide clues about the questions they would answer with ease and the ones they might try to avoid. Scrolling through the politician's Twitter feed, I stumbled on a photo. She was pictured arm in arm with leaders of La Resistencia, a far-right gang notorious for harassing opponents of Fujimorismo. The group's foot soldiers had recently cornered a soft-spoken, elderly prosecutor who had investigated Alberto Fujimori for extrajudicial killings years before, screaming "Communist! Terrorist!" as they closed in on him. In a country with a living memory of an internal armed conflict that claimed 80,000 lives, the episode didn't land lightly. I wouldn't be talking to just any Fujimorista, I realized. This was a hardliner.

That was challenging to reconcile with the way the politician greeted me, with a hug and a warm smile. When we got to the topic of my research, I chose my words carefully. "I'm here trying to understand whether justice has been served over the past few years in Peru, or not." I let the question hang in the air as I calculated what to say next. "But if I could start by asking a different question, what made you get involved with politics in the first place?" As her shoulders tensed and she inhaled, I realized the question, which I had hoped would make for a low-stakes icebreaker, had not landed that way. "I come from a region that had always been abandoned by the state," she began. She went on to describe how, due to lack of access to medical care, she had watched her younger brother die from a preventable disease. Thanks to Fujimori, she told me, a hospital was built in her town a few years later, and she never forgot it. The Fujimoristas, she believed, were the only politicians who had ever concerned themselves with extending the state

to the forgotten corners of Peru, and as soon as she could run for office with the party, she did.

This conversation was just the first of several in which the politician would open up to me about her experiences on the inside of Fujimorismo. To my surprise, the picture she painted was not a self-aggrandizing one. Instead, she candidly described the debates and turmoil on the inside of the party. Later, I was able to independently corroborate much of what she told me in interviews with other Popular Force leaders. The frankness of our conversation was due to the fact that I allowed the politician to believe I saw her as she saw herself—as a victim of injustice, not a perpetrator. I knew that not everything she told me was true: for instance, there was plenty of evidence to contradict her insistent claim that prosecutors had unfairly targeted the Fujimoristas or acted on a hidden political agenda. Still, as we talked, I noticed myself adamantly nodding along. At first, I dismissed it as a reflex. After all, I needed to keep the conversation going. But as time went on, I wasn't so sure. I started to wonder: how long can you act as if you believe someone without really believing them?

Lee Ann Fujii (2010) observed that interviews produce not only data, but also metadata: rumors, inventions, denials, evasions, and silences that add layers of meaning to interviewees' testimonies. In my experience doing fieldwork on how states sanction corruption in Peru, Colombia, and Guatemala, I have found that interviewees tend to subscribe to what could be called metanarratives. Metanarratives demarcate lines between truth and lies, good and bad actors, and means that do and do not justify political ends. More than mere collections of opinions, metanarratives are cohesive, overarching stories that give order to interviewees' observations and anecdotes. They supply the frames of reference interviewees use to make sense of day-to-day events. In polarized contexts, metanarratives can look less like contrasting perspectives, and more like two mutually incompatible pictures of reality.

In Peru, two metanarratives surfaced again and again, articulated by interviewees on either side of a polarized divide that emerged over the ongoing corruption investigations. On one side were critics of the prosecutors; a group that unsurprisingly included defendants in the corruption cases, but also their sympathizers in the judiciary, politics, and civil society. They believed prosecutors had trampled on due process to pull off high-profile arrests. Rather than investigating where there was probable cause, prosecutors had begun to investigate wherever there was potential for a media circus and the opportunity to drum up public support. The medicine—corruption investigations—had proven worse than the disease. Defendants undeniably spent

years in pretrial detention without facing formal charges, and the investigations had destabilized government after government. In order for the country to turn the corner, critics believed that the prosecutors must be reined in, even if it meant that some would continue to engage in corruption with impunity—though this last part went unstated.

Meanwhile, prosecutors, judges, and anti-corruption activists were committed to an entirely different metanarrative. This narrative rested on the belief that the prosecutions marked a turning point for democracy and the rule of law. For nearly all of Peru's 200-year history, the justice system treated elites and ordinary people differently, but now the corruption investigations were showing that no one was above the law. Prosecutors' methods were tough, but they had to be; they were investigating criminal organizations disguised as political parties. Defendants were dangerous; they had proven their ability to harass and intimidate judges and prosecutors—or worse. To end impunity and demonstrate the state was capable of administering justice, prosecutors needed to prioritize getting results, even if it meant pushing the law to its limits.

Over time, I came to realize that my interviewees did not just see each other's metanarratives as mistaken; they believed they were premised on lies. For prosecutors and their supporters, defendants' professed concern for due process rights was little more than a sham. Peru had always suffered from weak rule of law, but in decades past, citizens who suffered lengthy pretrial detentions and harsh sentences had been almost always poor or Indigenous. Now that elites found themselves on the receiving end of justice, they were merely deploying talking points about fundamental rights to mask their attempts to evade justice. Meanwhile, the other side insisted it was the prosecutors who were acting in bad faith: prosecutors and judges cynically portrayed themselves as embattled anti-corruption crusaders to gain the public's sympathy, but in reality they had come to hold all the cards. Anyone who questioned the anti-corruption campaign's merits was publicly branded an ally of impunity, which stifled debate. The judiciary wasn't strengthening checks and balances or the rule of law; it was tearing them apart.

Half the work of preparing for interviews—with the prosecutors, the hardline Fujimoristas, and everyone in between—was studying these metanarratives. When I wasn't with my sources, I was studying their metanarratives, watching the media they followed, and memorizing reference points from the news which I could deploy to show familiarity with both sides. Demonstrating familiarity with both metanarratives was critical for building contacts and conducting candid

interviews. When I managed to do it well, it allowed me to put my interviewees at ease. I was surprised by the misdeeds, strategic blunders, and self-critical reflections interviewees would share after I had gained their confidence.

Still, as my time in the field continued, something disorienting happened. Somehow, I found myself "buying into" both metanarratives at once. It became increasingly difficult to "play the part" because I was developing convictions of my own, informed by my interviews—some of which turned viscerally emotional. When I went to visit one politician under house arrest, she described in vivid detail spending years in prison on pretrial detention, racked by the anguish of not knowing whether prosecutors would press charges or simply let her go. Just a few days later, I interviewed a prosecutor who described the mysterious cars that trailed him through Lima and the pounding at his door that sometimes woke him at night. "If the Fujimoristas had won the last election, I would be in exile or in hiding right now," another prosecutor told me in a low whisper with fear in his eyes. Qualitative researchers know that evidence generated through interview-based fieldwork can motivate the revision of preconceived hypotheses. But the challenge I was dealing with was messier; there seemed to be truth to what sources on both sides of the divide told me, but their versions of even the basic facts were often entirely at odds.

Speaking fluently in both metanarratives wasn't easy before I became conflicted about who or what to believe. Afterwards, it became even more challenging. Sometimes, I stumbled. I was interviewing a prominent investigative journalist who had stridently supported the corruption prosecutions, and received threats for doing so, and our conversation was flowing. But when I asked a question about prosecutors' habit of leaking confidential information about their investigations to the press—a practice that technically broke the law—the journalist immediately froze up. Without intending to, I had let slip my familiarity with the metanarrative of the other side, and for the journalist, the result seemed to be disconcerting. "Of course, I can't discuss that," he snapped. The interview never recovered its openness.

At the same time, treading too lightly carried its own risks. When I refrained from questioning my interviewees' metanarratives, they could easily steer around difficult-to-discuss subjects. One night, I took an elevator to the Lima penthouse of a millionaire businessman who had long served as an advisor to the Fujimorista party. Midway through our conversation, I attempted to broach the topic of Popular Force's failure to stop the prosecutors. The businessman had been pacing back and forth across the room, but suddenly he stopped and

pointed to the notebook I was holding. "If you want me to tell you about our mistakes, you would need a few of those, because you would fill them all up." As much as I wanted him to go on, I froze. What if I seemed too eager? I let the silence continue for a few seconds too long, and to my disappointment, he moved on. The window of opportunity to dig deeper into the topic had closed. In other interviews I managed to walk the line without losing my balance, offering enough affirmation for my interviewees' metanarratives that they were later open to me questioning and challenging them.

When I arrived in Peru, it wasn't my first time doing extensive fieldwork in a challenging context. Having conducted several dozen interviews on vote buying and organized crime in Colombia, I felt confident that I could manage the challenges of fieldwork in a new country: building contacts, keeping calm in possibly tense situations, and quickly absorbing new information about the case on the ground. What I didn't expect was how difficult it would be to reckon with the opposing convictions I heard my sources express, which I absorbed by proxy. But in the end, this challenge proved to be the most generative part of my fieldwork. Immersing myself in both sides of Peru's anti-corruption divide confronted

me with a paradoxical reality of the case: prosecutors really had taken advantage of the weak due process rights baked into Peru's criminal justice system, but in large part that is also what had enabled them to finally hold elites accountable. Whether the rule of law was now stronger was difficult to say. No one was above the law, but now everyone was potentially subject to its abuses.

By the time I finished my fieldwork, I had learned to value the tensions that emerged from buying into conflicting metanarratives. I was less concerned with resolving these tensions, and more interested in the ways they might advance theory-building. In the future, researchers might use the concept of metanarratives to think about how they prepare for interviews, engage with interviewees, and use fieldwork to generate hypotheses. Interviewing sources and building contacts on either side of a polarizing divide requires walking a delicate path. Striving to understand my interviewees' metanarratives helped me meet this challenge, and might help future researchers balance affirming interviewees' preconceived beliefs with prompting them to further explain their own and others' action and behavior.

## References

Fujii, Lee Ann. 2010. "Shades of Truth and Lies: Interpreting Testimonies of War and Violence." *Journal of Peace Research* 47, no. 2 (March): 231-41. https://doi.org/10.1177/0022343309353097.

**QMMR** Qualitative & Multi-Method Research

# Notes from the Classroom

# Active Learning in the Qualitative Methods Classroom

Julia Lynch
*University of Pennsylvania*

It is the second week of the PhD semester in Research Design and Qualitative Methods for Political Science. A group of three PhD students is wandering through the building, looking for things to measure. They've been assigned to come up with a concept that is of potential political relevance, but also observable within the confines of our classroom building and nearby exterior spaces, and measurable within a span of 30 minutes' "fieldwork."

In preparation for this class session, the students read

about concepts and measurement from King, Keohane and Verba's *Designing Social Inquiry* (2021, 49-74, 150-68); Gary Goertz's *Social Science Concepts* (2020, Chapters 1-3); Giovanni Sartori (1970); and Richard Locke and Kathleen Thelen (1995). Each student also prepared a response to the following homework exercise: "Pick a concept that is used in political science that you think is particularly well operationalized and measured by at least one author. It could be very abstract (e.g., justice, regime), very concrete (e.g., occupation, campaign

advertisement), or something in between (e.g., social cleavage, political socialization, stability). Diagram the concept's dimensions, specify the measures, and discuss the intension/extension of the concept. What is it about the concept or its operationalization that makes it work well?" After this week, the exercises will be reviewed by a classmate, who provides comments[11]— but for the purposes of today's session, the homework exercise served as a warmup to prepare them for the in-class exercise. As a final preparatory step before the students went out to do their measurement, they were given twenty minutes to work with their groups in the classroom to decide on a concept and discuss their measurement strategy.

One group has decided to measure "security" in the context of our building, and it's going pretty well. They've been through the building counting fire alarms, keypads and remote cameras and making note of security-themed signage. They've observed the behavior of people entering the building. (Do they show their campus ID cards? Stop to identify themselves to the guard at the front desk?) And they've interviewed some members of the building staff to find out more about the security arrangements and procedures. Although there was initially some tension in the group about whether and how to combine counting with more qualitative assessments of security, things are going well. They're pretty sure they've nailed it.

They come back to the classroom, along with the other students who have similarly been sent off in search of things to measure, and get ready to report out their results to the rest of the class. I listen in on the groups preparing to present and observe that the "security" group had a particularly well-chosen concept and have interesting results to share. I choose this group to present first, aware that the quality of their work and their confidence in it will buoy them through what I know is coming next.

I ask the group to share with us the concept they chose, define the concept, tell us how they measured it and why they measured it that way, and then let us know what they found. Almost immediately, and despite the preparation described above, the wheels begin to come off the bus. I direct their presentation by asking questions that make it clear that the group's definition of security is internally undifferentiated (i.e., they have not defined its dimensions), and driven almost entirely by what they measured. Moreover, as the students complete their reporting out, it also becomes clear that they chose their measures based on what was available, and not because

they matched up with the domains that might be covered by dimensions of the concept.

Students begin to shift uncomfortably in their seats, in part in embarrassment for their colleagues and in part because it is dawning on them that they have done the same thing. Mindful of the need to spare the presenters from experiencing a sense of shame, I reassure them that this is the outcome the exercise was *designed to produce*. It happens every year, to nearly every group, because nobody is born knowing how to do this, and this is how we learn. Then I step up to the chalkboard.

I ask questions about the dimensionality of the concept of security, referring to Goertz's (2020) notions and diagrams. I scribble on the board, erase, redraw, etc., as the students begin to work through the idea that the concept has dimensions, and to identify which dimensions are relevant for the way that they might want to use it in the present context, or in a different context. Then I prompt them to relate those dimensions to the things that they have measured. Some of the measures match dimensions of the concept, while others don't. There are dimensions that lack measures, and dimensions that have more than one measure associated with them.

As they talk through what they did, and as other students in the class join in the conversation to offer suggestions or ask questions, the students in the "security" group become aware of things that they actually could have observed that would have matched the dimensions, as well as things they would have liked to measure but couldn't in that environment or given their resources (time, access to people and other sources of information). By the end of the debriefing session the students have a clearly diagrammed concept and some measures that match up with the dimensions of the concept, and they know where there are holes in their measurement. More importantly—for few of these students will go on to think about security in an academic building as a core theme of their research—the students have come out of the class session with a set of skills that I will watch them use repeatedly throughout the rest of their careers as PhD students.

This is a process I've repeated innumerable times in my 15+ years of teaching this seminar. In fact, each week of my research design and qualitative methods course is structured around some kind of active learning exercise. But this is one of my favorite exercises, for several reasons.

First, it embodies authentic learning: The exercise teaches students to perform initially unfamiliar tasks (concept diagramming and linking conceptualization

---

1 I have employed both peer and instructor grading for the weekly homework exercises, and each strategy has its pros and cons. I still use instructor grading when the class size is very small or when teaching online, but overall I prefer peer grading because it gives students a chance to reinforce what they have learned in class as they comment on their peers' exercises.

to measurement in a systematic way) that they will employ repeatedly (Herrington and Herrington 2005). The readings give them a language with which to link these processes to conversations in the discipline about method, which is also essential for their development as scholars. But the in-class exercise gives them a structured way to practice applying skills that they will use repeatedly in their own research (and in teaching, if they go on to an academic career).

A second reason I find this exercise valuable is that it delivers immediate impact. Some lessons about research design and methods take time and experience to truly resonate, and students may only come to see their value months or years later as they pursue their own independent research. This exercise, though, delivers instant bang for the buck. Students watch systematic concepts with well-chosen measures emerge in front of their very eyes on the blackboard as we talk through what each group has done. Early in the semester this is particularly valuable because it generates student buy-in; it gives students confidence that the active learning exercises they will confront throughout the semester, some of which will be uncomfortable, are worth doing.

Third, I love how this exercise helps to set the tone in the classroom early in the semester. Students need no specialized knowledge to do this exercise well. The readings provide all the background they need, and the homework exercise gives everyone a chance to prepare for what they will be doing in class. This means that a stats jock, a first-year PhD student who has never taken a class in political science, and a third-year student struggling with her prospectus because she's had no formal training in the basics of research design, are on a fairly level playing field as they go into the exercise, and can contribute equally useful insights. This promotes an inclusive classroom environment from the outset (Florian 2015). The fact that nearly all students so predictably "fail" the exercise, and then "succeed" as we work through their concepts and measurement on the board, also promotes a growth mindset that carries with students throughout the semester (and hopefully beyond) (Sahagun et al. 2021).

Finally, during the preparatory 20 minutes of the exercise, when students are working together in the classroom, it gives me the opportunity to observe group dynamics and interrupt unhelpful patterns early in the semester. For example, while teaching this class more than 10 years ago, I observed a group composed of one female and two male students preparing to measure a concept in a way that required them to identify the make

of cars passing on the street. I don't remember what the concept was, but I do recall that the female student initially took issue with it. When the male students continued without acknowledging her concerns, she tried again, explaining that she would not be able to identify the make of most cars. At her third unsuccessful attempt to be heard, I intervened: I identified what I had observed, and asked the group to find a different measure that they could all agree on. It was early in my teaching career, and I'm not sure that I handled the situation in the best possible way—but I knew that the gender dynamic that the exercise had revealed needed to be nipped in the bud. Some years later, in an unusually small class that had only one female student in it, I noticed that that she had begun taking notes for her group as they worked on this exercise. I stopped them to ask how they had decided who should take notes, and as I had suspected, it emerged that they had not discussed it; the female student had simply started to do it. This prompted a very productive discussion about gender norms and assumptions in group work that several of the (male and female) students in the class have told me stayed with them.

*

I began teaching research design and qualitative methods using active pedagogy—learning by doing—well before I knew that there was such a thing as active pedagogy or a body of research showing its efficacy (Talbert and Mor-Avi 2019). I have always learned methods most thoroughly when I am trying to apply them, so it seemed to me like a natural way to teach methods. To be sure, short papers also provide important opportunities to learn through application, and this course requires three 10-page papers in addition to the weekly assignments. But active pedagogy provides the ability to give and receive immediate feedback when students are first learning to use a new skill, as the example I've presented illustrates. Active in-class exercises take time to prepare and hone, but at this point I can't imagine teaching research design and qualitative methods effectively without them.

My syllabus for Research Design and Qualitative Methods is available on my website: https://web.sas.upenn.edu/jflynch/teaching/. The syllabus includes the homework exercises for each week, but not the in-class exercises. Please feel free to email me at jflynch@sas.upenn.edu if you are interested in other examples of in-class exercises.

# References

Florian, Lani. 2015. "Conceptualising Inclusive Pedagogy: The Inclusive Pedagogical Approach in Action." In *Inclusive pedagogy across the curriculum (International Perspectives on Inclusive Education, Vol. 7)*, edited by Joanne M. Deppeler, Tim Loreman, Ron Smith, and Lani Florian, 11-24. Bingley: Emerald Group Publishing Limited.

Goertz, Gary. 2020. *Social Science Concepts and Measurement: New and Completely Revised Edition*. Princeton: Princeton University Press.

Herrington, Tony, and Jan Herrington, eds. 2005. *Authentic Learning Environments in Higher Education*. Hershey: IGI Global.

King, Gary, Robert O. Keohane, and Sidney Verba. 2021. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. rev. ed. Princeton: Princeton University Press.

Locke, Richard M., and Kathleen Thelen. 1995. "Apples and Oranges Revisited: Contextualized Comparisons and the Study of Comparative Labor Politics." *Politics & Society* 23, no. 3 (September): 337-67. https://doi.org/10.1177/0032329295023003004

Sartori, Giovanni. 1970. "Concept Misformation in Comparative Politics." *American Political Science Review* 64, no. 4 (December): 1033-46. https://doi.org/10.2307/1958356

Sahagun, Miguel A., Randy Moser, Joseph Shomaker, and Jenna Fortier. 2021. "Developing a Growth-Mindset Pedagogy for Higher Education and Testing its Efficacy." *Social Sciences & Humanities Open* 4 (1): 100-168 https://doi.org/10.1016/j.ssaho.2021.100168.

Talbert, Robert, and Anat Mor-Avi. 2019. "A Space for Learning: An Analysis of Research on Active Learning Spaces." *Heliyon* 5, no. 12 (December): e02967. https://doi.org/10.1016/j.heliyon.2019.e02967.

QMMR

Qualitative &
Multi-Method
Research