ANALYTICAL ESSAY

The Politics of Data Access in Studying Violence across Methodological Boundaries: What We Can Learn from Each Other?

NOELLE K. BRIGDEN

Marquette University

AND

ANITA R. GOHDES

Hertie School

In this article, we investigate where the ethics of data collection and access of two widely disparate methodological approaches studying violence intersect, and we explore how these respective intellectual communities can learn from each other. We compare and contrast the research strategies and dilemmas confronted by researchers using quantitative methods to collect and analyze "big data" and those by researchers conducting interpretivist ethnography grounded in the method of participant observation. The shared context of participant vulnerability produces overlapping concerns about our work. With shifts in quantitative conflict research to examine the microdynamics of violence, quandaries of confidentiality and the ethics of exposure have become increasingly salient. At the same time, ethical dilemmas that arise in the large-scale collection of data offer important points of reflection regarding the ethics of participant observation as it is performed in ethnographic research. Instead of focusing on areas of disagreement, we suggest that interpretivist fieldworkers and quantitative researchers can learn from how the politics of information materialize across divergent research methods.

Keywords: data access, ethics, violence

Introduction

Against the backdrop of global surveillance disclosures, scandals of social media platform abuse in the context of elections, and rapid developments in the employment of technology for social control, societies are awakening to the power of information and the politics of privacy in the twenty-first century. Governments exploit new, invasive forms of controlling their own populations, powered by vast amounts of data collected through facial recognition tools, social media content, and metadata. Under the auspices of the *War on Terror*, democratic regimes have granted themselves new powers to access and use private communications in the name of

national security. Private sector actors with an extensive reach into the everyday habits of citizens have commercialized personal information, creating new markets for data once considered private. Privacy and civil rights advocates campaign for new protections, but legal systems have been slow to adjust to technological change and the rising power of corporate data gathering. In response, some observers have heralded "the death of privacy" (Preston 2014).

Researchers across a spectrum of disciplines are also awakening to the politics of data collection, accessibility, and protection (Boyd and Crawford 2012; Criado-Perez 2019). In reaction to an outcry over the need to reestablish academic credibility in a "post-factual" political era, scholars have redoubled their efforts to generate norms surrounding data access, transparency, and the replicability of research findings. The Data Access and Research Transparency (DA-RT) initiative led by the American Political Science Association (APSA) aims to generate discipline-wide standards for transparency at peer reviewed journals and is the central effort to achieve these ends.[1] The initiative has been endorsed by many of the discipline's leading journals, and has also found considerable support among scholars based outside of the United States. However, the initiative has not been without controversy and has renewed the debate over the methodological and ethical dimensions of data collection and dissemination (Parkinson and Wood 2015; Pachirat 2018; Thaler 2019).

Many of the ethical considerations guiding these discussions are geared toward addressing systemic issues of how to establish legitimacy in social science. The legitimacy of scientific claims is hereby tied to results being a "product of publicly described processes that in turn are based on a stable and shared set of beliefs about how knowledge is produced" (Lupia and Elman 2014, 20). Importantly, the agreed upon processes and beliefs vary widely across research communities (Büthe and Jacobs 2015). For positivist social science traditions, the key indicator for such credibility has been reproducibility through openness (Lupia and Elman 2014). For interpretivists, it has been reflexivity, defined as the development of a self-conscious, systematic, and explicitly stated awareness of the relationships between the researcher, the participants, the research findings, and power (Parkinson and Wood 2015; Krystalli 2018; Pachirat 2018; Thaler 2019). The ongoing discussion within political science regarding shared norms on data access and transparency has therefore highlighted that researchers with different methodological and epistemological positions tend to disagree about the practices, protocols, purposes, and even the very meaning of transparency (Büthe and Jacobs 2015).[2]

In this article, we focus on political science research in the context of violence in order to investigate where the ethics of data collection and access of two widely disparate methodological approaches intersect and how their respective intellectual communities can learn from each other.[3] We focus on our own scholarly communities—namely, research using quantitative methods to collect and analyze large amounts of data and interpretivist ethnography grounded in the method of

---

[1] See the 2012 revision of the APSA's *A Guide to Professional Ethics in Political Science* and the 2014 statement for detailed requirements being adopted by many top journals in the discipline.

[2] For the purposes of this conversation, we adopt a broad understanding of transparency, defined here as a principled commitment to "providing a clear and reliable account of the sources and content of the ideas and information on which a scholar has drawn in conducting her research, as well as a clear and explicit account of how she had gone about the analysis to arrive at the inferences and conclusions presented—and supplying this account as part of (or directly linked to) any scholarly research publication" (Büthe and Jacobs 2015, 2). This definition leaves the debate over the specific protocols, practices, and standards for transparency open to debate.

[3] We focus on political violence because many of the ethical issues we address here become immediately obvious to this field of research and because this is the field we both work in. The issues we address here, however, are likely to carry into other areas of research that deal with the repercussions of political power struggles and human vulnerability. For example, big data and interpretivist work in health, demography, job security, and disability, to name just a few, are likely to provoke comparable ethical concerns.

participant observation.[4] From these divergent methodological positions within the discipline of political science, both of us work on themes broadly related to state and nonstate violence, criminalization, and clandestine resistance to a variety of human rights violations. The shared context of participant vulnerability produces overlapping concerns about our work. Instead of focusing on areas of disagreement, we aim to explore the many ways in which the safety and ethical dilemmas experienced by qualitative researchers engaged in fieldwork with vulnerable populations have analogue for quantitative researchers engaged with individual-level data about vulnerable populations, and vice versa.

We discuss how the multiplying array of data sources available to study these sensitive issues remotely, paired with increasingly blurred boundaries between private and public worlds online are confronting quantitative researchers with dilemmas that had once been understood as particular to the domain of qualitative fieldwork. With shifts in quantitative conflict research to examine the micro-dynamics of violence, these quandaries of confidentiality and the ethics of exposure have become increasingly salient. At the same time, dilemmas that arise in the large-scale collection of data offer important points of reflection regarding the ethics of participant observation as it is performed in ethnographic research. We suggest that interpretivist fieldworkers and quantitative researchers can learn from one another by more closely discussing how the politics of information similarly materialize in divergent research methods.

What Information Are We Talking About? "Data" Collection and Access in Political Violence Research

Quantitative Data Collection

Traditionally, much of the quantitative research on political violence has been concerned with the collection, description, and analysis of state characteristics, including both domestic traits and external relationships with other countries (e.g., Oneal and Russett 2001; Fearon and Laitin 2003; Harff 2003; Leeds 2003). Quantifiable state characteristics would include indicators such as military spending, economic prosperity, or alliance membership. Comparative coding of state characteristics has traditionally placed the emphasis on aggregate information that would conceal individual-level characteristics, behaviors, and experiences. Examples include descriptions of regime type, as well as standards-based measures of conflict, repression, and interstate relations. Over the past decades, quantitative research on political violence has witnessed a pronounced move toward studying the characteristics of violent events, episodes, and actors at the subnational level (see e.g., Eck and Hultman 2007; Cederman and Gleditsch 2009; Lyall et al. 2013). This development has been accompanied by more fine-grained collection of both experimental and observational data. Whereas aggregate state indicators generally avoided the disclosure on individual-level information, disaggregated analyses of violent contexts frequently rely on information related to individual actors to infer broader dynamics of conflict. With the spread of digital communication technology across conflict zones and the accompanying availability of massive amounts of data, quantitative researchers on political violence are now working with progressively more granular information sources (e.g., Shapiro and Weidmann 2015; Mitts 2019; Gohdes, forthcoming). Statistical analyses at an increasingly disaggregated level with the help of innovative

---

[4] It is worth noting that one of us has been trained and is based primarily in Europe and the other in United States. In these institutional contexts, there are divergent cultures, laws, ethics protocols, and bureaucratic practices that guide research. For example, in the United States, the Institutional Review Board (IRB) is a mandated process at universities. In Europe, the European Research Council has specific guidelines on reviewing ethics for projects it funds (see e.g., European Commission 2019). At the national level, the existence, involvement, and requirements with regard to ethics reviews vary widely.

data collection techniques have brought with them a wealth of new insights on the causes and consequences of political violence.

Examples of disaggregated data that have been collected through both primary and secondary research may include individual- and group-level information that specifically recounts the characteristics of committed or experienced violence, including identifying information and political preferences of victims, bystanders, and perpetrators. To illustrate, a recent study by Gohdes (forthcoming) uses individual-level data on more than sixty thousand victims who were killed by the Syrian regime. The original data were collected by a variety of documentation groups active within Syria, and cleaned, canonicalized, and linked by Gohdes and coauthors (see Price, Gohdes, and Ball 2014, 2016). The data include information on each victim's name, location, and date of death, as well as a description of the circumstances within which the person was killed. Many of the records include weblinks to pictures, media depictions, or videos as they were posted to various social media platforms. The speed, digital cross-linkage, and level of detail recorded on the fate of each individual victim would not have been possible without both the courageous work of documentation groups *and* the convenience of modern communication and information processing technology.

Less directly, data collection efforts may touch on information that gives insight into individual- or community-level socioeconomic, ethnic, religious, and demographic characteristics. Digital data collection may include the scraping of content from online government sources, news websites, or social media platforms, including both "real-time" and longitudinal data. Digital trace data may include metadata such as phone call records or other forms of human actions and interactions that leave a digital footprint. Finally, advances in data processing technology have vastly facilitated the digitization of historical material from archives, thus making it accessible to a significantly wider audience.

The means of data collection by amassing large amounts of fine-grained information, including the employment of new survey measurement techniques, have grown more technically sophisticated in nature. Precisely because these new methods and techniques allow for a more fine-grained testing of theoretical mechanisms, the associated data sources bring with them dangers of harm that may negatively impact both targeted and "unintentional" research participants. An obvious example is the gathering and analysis of user-generated social media data, such as those obtained through the Twitter API. On the production of social media data during crises, Crawford and Finn (2015, 496) write "In a crisis, someone may be reporting what they see in a 'citizen journalism' style, while also alerting friends and relatives to their wellbeing, while also recirculating both verified and unverified reports of others." Correspondingly, social media posts created in the context of political violence may constitute a mix of reporting, personal communication, rumors, and entirely unrelated content. Even though social media users are required to agree with the platform's terms of service, it is not immediately clear whether they would knowingly consent to their content being lifted out of the context of extreme political unrest and be reposted and analyzed in the context of a research study. As we will discuss further below, the immediate and long-term risks associated with the reposting and republishing of situational content are often hard to gauge for both researchers and research participants.

In investigating these ethical dilemmas, we see important parallels to those found in interpretivist research that primarily relies on participant observation. As we explore further below, granular quantitative data dealing with violent contexts may expose information about research participants, thus leaving them vulnerable to dangers in similar ways to what an interpretivist's field notes would. The next section introduces the use of participant observation and the creation of field notes in the context of research on violence.

Interpretivist "Data" Collection: Participant Observation and the Role of Field Notes

In tandem with the shift by quantitative scholars of political violence toward disentangling fine-grained individual level data, qualitative political scientists have also recently embarked on ethnographic fieldwork in conflict zones, clandestine spaces, and authoritarian regimes, shifting our analytical focus to the micromechanisms of these processes (e.g., Wedeen 1999, 2008; Wood 2003; Fuji 2009; Autesserre 2010; Pachirat 2013; Parkinson 2013; Durán-Martínez 2018; Brigden 2018a; Idler 2019). Fieldwork on violence raises urgent ethical concerns, which puts many of these political scientists at odds with recent disciplinary attempts to adopt standardized transparency norms (Parkinson and Wood 2015).

Ethnographic fieldwork is grounded in the method of participant observation, aimed at understanding the internal logic of alternative worldviews (Schatz 2009, 5). Participant observation, at its most basic level, can be understood as a careful contemplation of the life experience and daily practices of another person or people and requires firsthand immersion in a social context to be achieved (Brigden and Mainwaring 2019; Schatz 2009). Importantly, many discoveries made during this process cannot be traced to structured interviews, which can be illustrated by a straightforward interview methods table, as one would want to include in an appendix (e.g., Bleich and Pekkanen 2015, 11). Participant observation is usually captured in the researcher's field notes, which include descriptions of a wide range of interpersonal interactions, researcher experiences, visual observations, and even emotional or sensory reactions to those experiences. They follow a carefully structured method of description to discipline thought and evoke discoveries, which distinguishes the practice from journal writing (Emerson, Fretz, and Shaw 1995).

As such, field notes from participant observation are an ongoing contemplation from which inductive insights may be drawn; they are a systematic method of thinking about the social context in which the researcher is immersed. Findings, field notes, and observational process cannot be disentangled (Emerson, Fretz, and Shaw 1995, 11). Such experiential and inductive moments of discovery within field note taking are therefore not "data" in the traditional sense of understanding and thus cannot be treated as perfectly analogous to data collected for quantitative analysis. Instead, tacit knowledge from immersion in a social context and a trained, self-conscious sensibility for judging field notes and making discoveries within them is a way to evaluate evidence. Interpretivists must contextualize their work in the particular setting of the field to be convincing, rather than justify sampling as representative to a broader population (Cramer 2015, 18; Pachirat 2018, 149). Field notes are written with the researcher herself in mind; while rereading them many unrecorded details may flood back to her, ranging from smells to sounds and unspoken impressions, providing a further context for her interpretation of events (see Cramer 2015). The expectation that notes would be made public, particularly in violent contexts, would change the nature of that process dramatically and disrupt the relationship of trust between participant, researcher, and interpretive process (Parkinson and Wood 2015, 24). The researcher does not simply analyze what is written in their notes as objective fact but instead draws on a richer experiential context, sorting through it using a reflexive sensibility, where the field notes serve as device to propel critical thought. Participant observers record even the most mundane, fine-grained details, some of them deeply personal, but do not necessarily treat such records as facts or data.

An excerpt from Brigden's notes from fieldwork conducted along the clandestine migration route across Mexico in 2010 provides an example of this style of note taking. The researcher describes scenes without yet knowing their analytical value, but in the process of writing those descriptions, she discovers new insights. In this scene, a young man had been thrown from the freight train that Central American migrants ride north. In the migrant shelter office, Brigden peered over

the shoulders of other shelter staff at the images of his dead body and his personal effects. Local police authorities had asked the shelter to help identify him. The migrant shelter staff called some of the boys that hang around the train tracks in to identify the corpse in the photos. The following passage comes from Brigden's notes (recorded by her on November 6, 2010), describing the rituals and drama of this death, as she played a dual role as shelter volunteer and researcher:

> Across the yard [of the migrant shelter], I notice XXX hugging a young man. He has a broad pale face, a bright white shirt and new white sneakers. A baseball cap sits askew. He appears to be one of her boyfriends. They hug, and she cries happy tears. She reads a piece of paper that he gave to her. I recognize the boy from the night on the tracks when I went with Padre to pass out bread [and met gang members huddled cold and hungry along the train tracks].
>
> Inside the office, another young man dressed similarly looks at the photo of the dead boy. He is the boy with the tear tattoos from a few days prior. His girlfriend bursts into tears when looking at the photo. His face, too, seems to grieve. His girlfriend turns to him and buries her face in his chest. He holds her. I leave. . . .
>
> Moments later, Two Tears and his girlfriend emerge from the office. He walks with his arm looped over the shoulders of his girlfriend with a confident, slow, swinging cadence, almost as if he cannot be bothered to walk or even stand. He struts. It's cool. The girlfriend chews gum, make up and face intact. I have a hard time reconciling this sauntering image with the nervous, compassionate boy in the office.

Brigden's capacity to watch the transformation in their demeanor inside and outside the office proved to be a key to understanding survival strategies along a violent migration route. The process of writing down the scenes observed during daily activities forced her to reckon with these subtle changes, and it alerted her to how the performances of multiple identities shifted with social context and audience. If she had not been taking the most detailed and descriptive notes possible about as many interactions as possible throughout the day, she perhaps would never have noticed the fluidity in the performance of masculinity she had witnessed, as "Two Tears" moves from what Goffman calls "Back Stage" to "Front Stage" (Goffman 1959). Brigden began to understand improvised performances as a necessary means to navigate violence and uncertainty, and she began to reflect on her own performances as researcher, shelter staff, friend, mother, foreigner, and woman, as she too grappled with this social setting. This insight about improvisations on social roles and identities, which she dubbed "survival plays," became a central finding in her study, and such observations also helped her better interpret and contextualize information provided more directly in interviews.

Ethical Dilemmas in Data Sharing—Where Big Data Analysis and Ethnography Intersect

We discuss the ways in which data access and collection may endanger both targeted and unintentional research participants in oftentimes unanticipated ways and how these dangers can be found in both interpretivist and large-scale quantitative research. "Targeted research participants" are those individuals explicitly examined to learn about social and political practices during a research project.[5] These may be specific individuals—for example, selected through survey sampling. Or they

---

[5] These participants have presumably given their informed consent to participate, though, as the excerpt from Brigden's field notes makes clear, some methods (such as participant observation) create difficulty in negotiating this consent. Furthermore, there is no expectation that all the stakeholders, including what we call "unintended participants," have given their informed consent to participate. For example, an adult father's decision to be interviewed may impact the safety of his wife and children, but those dependents are not generally asked to consent to his participation in research. A migrant may give information about her own survival strategy to an interviewer, and the visibility of this information might undermine the survival strategies of other migrants, but protocols of informed consent generally

Table 1. Categorizing risks of data sharing for research participants

	Immediate	Long-term
Individual	- Failure to gauge current power relations, exacerbating existing vulnerabilities to violence	- *Geopolitical Uncertainty*: Shifts in power dynamics create new vulnerabilities for individuals who previously felt comfortable sharing/publishing information
	- Identification of individual, i.e., failure of confidentiality or anonymity	- *Technological Uncertainty*: New technologies and data collection complicate de-identification
Collective	- Harmful policy decisions that may impact stakeholders beyond targeted participants	- Information is lifted out of context
	- Existing data regulations inadequate for "new data"	- Revealing survival strategies of an individual renders those strategies potentially ineffective for the people that follow

may be specific groups, such as those who made references online about a certain violent event using a designated hashtag or those who witnessed or were involved in a specific violent episode. "Unintended research participants" are the broader stakeholders: individuals or groups that are not directly observed in a specific study but may, nevertheless, be adversely affected by the research process, the research publication, or the public accessibility of collected data. Simply by having a stake in the information, they are de facto participating in a study that they might know nothing about. Entire communities might suffer "collateral damage" from research that makes sensitive information visible to a new audience, even when none of the members of that community directly participate in the observations. We focus here on the ways in which data sharing in the aftermath of research activities may endanger either of these groups.

We distinguish temporally between immediate and long-term dangers associated with the sharing of data. As explained by Thaler (2019), "any geographic and social setting in which research was conducted will change, in both subtle and overt ways, after the researcher has left the field. These temporal changes make it difficult to anticipate potential downstream risks of research conduct and participation, rendering 'transparent' sharing of data and research procedures problematic." These dangers can affect individual research participants or have more collective implications. Table 1 summarizes these risks as they can pertain to both quantitative data and information collected from participant observation.

Making information that was collected for a research project on violence accessible to third parties can pose immediate dangers to individual research participants. Such information—whether coded in a quantitative way or captured in prose—runs the risk of identifying participants and revealing characteristics about them that were previously *not public or not easily accessible*. What both granular "big data" and field notes have in common here is that they tend to provide a certain degree of contextual information that can increase the risks of identifiability in ways that aggregate data previously seldom did. Even without the mention of names or addresses, field notes may provide context on an individual's living conditions, social relations, or migration histories that may allow adversarial actors with additional hidden knowledge to re-identify individuals and exploit the revealing of their preferences, experiences, or actions in harmful ways (Parkinson and Wood 2015; Krystalli 2018).

only recognize the targeted participant's right to decide, not a larger process of collective decision-making and responsibility. In the case of digital content data, users of social media platforms may have accepted the terms of service of a website when signing up to use it, which frequently include clauses about data usage, but may not be fully aware of the extent to which their content will be shared and analyzed.

Similarly, the assembly of multiple sources of granular data as they may have been collected for quantitative analyses can facilitate the unique identification of individuals, even if generic personal information has been removed (Narayanan and Shmatikov 2009).

In the cases of both field notes and data collected, public access can endanger individuals, even if the information they entail has already been publicly available or constitutes common knowledge. Providing access to accumulated information— either as it is included in the written observations of a field researcher or in the form of merged databases—can bring to attention previously hidden connections, relationships, histories, and contexts that risk having harmful personal or political effects for research participants. This increased attention generated by scholarship potentially creates new political incentives for retaliation against participants, even if the information had been previously publicly available. The involvement of new audiences generates new political meanings for old information and potentially provokes violent reactions, complicating the ethics of making clandestine practices visible in research, activism, or journalism (Brigden 2018b).

Importantly, legal scholars have argued forcefully that there are clear structural barriers that make it unreasonable to assume that individuals can self-manage their own privacy in the age of mass data collection and digital surveillance (Solove 2012, 1881). Besides individuals' preferences for privacy that need not be not related to immediate personal risks, researchers frequently fail to fully grasp local and more global power dynamics that can have adversarial repercussions. Individual victims or perpetrators may be persecuted by armed actors for revealing noncompliant preferences or past loyalties that are at odds with current allegiances. Dangers related to the publication of field notes or quantitative data may not always lead to physical harm but may instead lead to loss of social status or reputational damage for research participants who share vulnerable information, either while in direct interaction with the field researcher or indirectly in the information later collected by the quantitative researcher. Research participants' histories may be exposed to their family or wider community in ways that make them vulnerable to exclusion or ostracization.

Access to such information can also be collectively harmful for individuals and groups well beyond the targeted participants, despite the fact that researchers' informed consent procedures and protections generally focus solely on the targeted participants. In the example provided in Brigden's field notes given above, she had the express permission of the shelter to record any of her experiences or conversations, and she had even been granted unlimited access to their own database of information about migrants, suspected smugglers, and human rights violations, for the purposes of her research. She had also been approved to conduct participant observation by her university Institutional Review Board, a legally mandated ethical review process in the United States, but only semistructured interviews required a formal informed consent procedure.[6] Thus, she had a right to be in the office, observing as well as acting as a staff member. Nevertheless, "Two Tears" and his girlfriend were grieving privately, and they had not consented to be part of the study. Brigden recorded many of these intimate observations with an expectation that she would mediate access to the field notes, since she neither felt comfortable with the quality of consent under such circumstances nor did she want to reveal some situations in which knowledge of "back stage" performances would render them obsolete as survival tools. Survival plays need a back stage to be compelling. The stakes here go far beyond risks to anonymity and confidentiality because entire communities

---

[6] Presumably, field notes from participant observation were not collected for the purpose of generalizable scientific knowledge, not used as data, and therefore, not subject to this bureaucratic legal process. Brigden does not consider her field notes to be data.

depend on the viability of such strategies. Thus, even though this situation would "only" involve the visibility of two individuals, unmediated exposure could cause harm to an entire community of people.

Contexts and implications based on field notes or "big data" may also form the basis of harmful policy decisions; armed state or nonstate affiliated actors may exploit intimate case knowledge about the whereabouts and methods of resistance of vulnerable populations, and government agencies may benefit from accessing data and analysis to design repressive responses. In some instances, government-funded research projects may be compelled to share their data with government agencies (such as intelligence services), which may be directly harmful for the research participants and their extended communities.[7] Indeed, the heated nature of the controversy over the attempt to establish disciplinary-wide standards for research transparency stems from the power exerted by financial, publication, and policy gatekeepers who can mandate them and the potential impact of these mandates on resources for researchers with legitimate ethical or methodological objections to unmediated data sharing. Funders, both government and private sector, sometimes mandate data sharing, putting this subset of researchers in an uncomfortable position to negotiate (e.g., Krystalli 2018). At a global political moment when governments' commitments to human rights have wavered and refugee populations face increasing security scrutiny, the intersections of state power and funding mandates for research transparency become ever more troubling.

In many ways, the rapid expansion of "big data" that can be collected and analyzed presents an even more precarious challenge to the traditional issues raised by qualitative contextual information. The publication of such data may bring with it immediate and unforeseen consequences, in part because regulations pertaining to the lawfulness of collecting and distributing such data tend to lag behind their employment in research and development and tend to vary widely across domestic legal landscapes.[8] Because regulations pertaining to data access are unlikely to be able to keep up with changes in data production and availability, we contend that scholars have an ethical obligation to consider the harmfulness of their data collection and accessibility efforts, regardless of the legal context they are operating in.

There are also important long-term dangers that researchers within both research traditions under investigation here need to consider before making the information they have collected accessible to a broader public. Researchers may be cognizant of and sensitive to the dangers of publishing information, given current power dynamics, but those dynamics may shift in unexpected ways, thereby exposing individuals to harmful repercussions at later stages (see Parkinson and Wood 2015; Thaler 2019). A prominent example for this was the oral history project established by the Boston College's Burns Library in 2001 (known colloquially as the "Boston Tapes"), which recorded over two hundred interviews with nonstate armed groups involved in the Northern Ireland conflict (Breen-Smyth 2019). Although the archives were intended to be kept under embargo until the interviewees had passed away, a book referring to the material was published following the death of two key research participants, and, subsequently, Northern Irish authorities requested full access to the archives, thereby endangering not only the remaining research participants but also all of those implicated in the recordings, as well as their broader communities. More generally, fine-grained data on ethnicity, religion, education, political preferences, or economic status may only become controversial after political upheaval,

---

[7] We thank a reviewer for highlighting this practice.

[8] Research in computer science, in particular cryptography, has actively been working on solutions for issues related to sharing sensitive data, including extensive research on differential privacy (see e.g., Dinur and Nissim [2003] and Dwork [2011]).

newly developed polarization, or economic decline. Individual characteristics were previously seen as uncontentious but may then suddenly run the risk of becoming the basis of discrimination, social exclusion, or repression. In addition, data that was originally deemed "unidentifiable" or harmless may become quite sensitive and potentially dangerous when combined with data collected at a later stage (Solove 2012, 1889).[9]

Information that is made publicly available can also risk taking on a life of its own. Data may be lifted out of context and used in ways it was not intended to be used (Murgia 2019). This is true for information gleaned from participant observation but also for "big data" that was collected with a specific usage in mind and then applied to a different setting, where it is either abused or used to draw false conclusions. We do not wish to imply that it is researchers who collected the original information who should bear responsibility for this, and we hope to avoid blaming individuals for such unintended consequences, when complex power structures and contingencies condition these outcomes. Nevertheless, reflections on the possible abuse of one's research should form a central part in deliberations about the ethics of data access in any project. We also do not wish to tie the hands of researchers or limit the scope of inquiry. However, the contingencies of power and knowledge require that researchers grapple with the possibility that some discoveries might simply be too dangerous.

This discussion of short- and long-term risks pertaining to information access in research on violence demonstrates the enormous challenges related to gauging the present and future harmfulness of making one's research process more open. The dilemmas presented here have no easy, straightforward, or "one-size-fits-all" answers. Risks and dangers are highly context dependent and require a substantive investment on the side of the researchers. However, we contend that precisely because these questions transcend methodological boundaries, it is crucial for researchers to acknowledge that important work on ethical standards has been developed in both ethnographic and "big data" research. For example, the Responsible Data Community Platform has been working on definitions and standards that ethical rights-focused data collection, compilation, analysis, and publication should prioritize (Responsible Data 2019). They argue that researchers need to consider the *unknown unknowns*: "we can't see into the future, but we can build in checks and balances to alert us if something unexpected is happening." Analogously, but in the interpretivist tradition, Brigden and Hallett (2020) discuss the ethical obligations that ethnographers studying violent contexts have, arguing that they "must constantly consider past, present and future consequences," including "situation[s] in which the past is not necessarily a guide to the future." Indeed, Thaler (2019) convincingly argues that for both positivists and interpretivists, conducting research on *violence*, in particular, these issues of positionality, contingencies in data production, and uncertainty over time have heightened importance.

We argue that bringing together these discussions and standards in a problem-oriented way that recognizes progress made within a different methodological community will help advance safe and ethical research standards without putting research participants, whether targeted by or collateral to the investigation, at risk. In the following, we present two examples of areas in which we believe our respective research communities can learn from each other.

---

[9]The fact that most "big data" collection necessarily relies on information processing tools for the collection and analysis of information render the ethics surrounding the re-identification and abuse of data particularly difficult. An example of this would be where the employed tools are built and provided by corporations that cooperate with law enforcement (see e.g., Glaser 2019).

What Can We Learn from Each Other?

Taking Positionality Seriously: Implications for Big Data Research

Ethnographers have developed key protocols, strategies, and shared standards for evaluating the quality and ethics of their empirical and analytical output. These include the practice of careful coherent theorizing about the relationship between the researcher's own identity, interactions with participants, and the material and social practices that delineate the boundaries of the "field" and power structures. It may prove helpful for quantitative social scientists to contemplate these protocols, strategies, and shared standards rooted in ideas of reflexivity and positionality. In the interpretivist tradition, credibility of research is oftentimes established through reflexive practice, which requires researchers' explicit exploration of their relationships to participants and the power structures that shape that relation. It necessitates sustained engagement with concepts, the social context under study, and field observations over a period of many years (see e.g., Bourgois 2001).

The goal of interpretivist ethnographic intellectual work can therefore be understood as a continual *process*, an ongoing analytical struggle and (self)criticism without finality, and not necessarily knowledge understood as a final *product*, not a "discovered" object or accumulated facts. Similarly, its ethics are best understood as a perennial contemplative grappling and self-awareness. The goal is the process itself, and the moment that ethics or knowledge become accepted uncritically, above struggle, the process has failed. This process, as pointed out by Pachirat (2018, 149) "encourages reflexivity about positionality and an examination of power . . . This reflexivity extends as well to the potential impacts and effects of the politically and socially legitimated 'knowledge' produced through the researcher's embodied interactions with the social world." The embrace of this understanding of knowledge and ethics as open-ended process generates a set of strategies and shared standards that may prove instructive for quantitative researchers working with new, large amounts of data to avoid harmful and potentially unethical data collection and publication.

This tradition of reflexivity and thick transparency emerged from a process of disciplinary reckoning of the historical ties between the discipline of anthropology, ethnographic method, racism, and imperialism (Gough 1968). Indeed, in anthropology, this roiling process is still ongoing and remains deeply contentious, as new debates erupt over the appropriate boundaries between militarism, the state, funding, and ethnographic research (e.g., AAA Executive Board 2007). A more public and thoughtful reckoning of the relationships between both qualitative and quantitative political science, funding streams, data access/sources, and government projects of social control would enrich the developments sought by transparency advocates in our own discipline.

In quantitative research, the positionality of the researcher may seem less obvious at first glance than it is in the written recollections of participant observation. Yet there are many important and oftentimes critical ways in which data collection is heavily influenced by the researcher's own identity. At an institutional level, a researcher's access to resources may principally affect the type of data they are able to gather. The availability of both financial and technical research support will determine the quality, quantity, and breadth of data available for a research project. Beyond institutional factors, a researcher's own identity is likely to influence the types of questions they might ask and consequently affect the types of data they will ultimately seek to collect and analyze. This includes a researcher's training, disciplinary background, gender, geographic location and upbringing, and political preferences, as well as the political climate in which they predominately work and many factors that may be more subtle in their effect. While replicability and transparency of the research processes in quantitative research may be more straightforward then they are in interpretivist work, the question arises as to what extent

researchers are able to reflect on their position in determining what data *were not included* in their collection efforts, either because they didn't deem it relevant or because it wasn't available to them. In sum, we contend that critically investigating without prejudice if and how these factors may influence both the questions asked and types of answers expected to be drawn from the data is a useful practice for researchers studying violence from all methodological angles.

Finally, quantitative researchers are likely to benefit from questioning to what extent their research objectives and selection of research participants is determined by the status quo of available information (see also Weidmann 2013; Price and Ball 2014). With an exponential increase in data revolutionizing much of the quantitative social sciences, reflection on how research questions are formed by what data are available to researchers and the implications this has for the ethics of data sharing and access has become more important than ever. Furthermore, acknowledging and openly discussing the inductive nature of research processes that oftentimes get concealed as projects evolve and take on a life of their own would constitute an important part in the reflections of both "big data" and interpretivists research. This is certainly not to posit that research questions borne out of the availability of specific sources of information or contexts are per se problematic; quite to the contrary. As Duneier puts it: "In much of social science, especially much of quantitative research using large data sets, a research design often emerges after data have been collected . . . like quantitative researchers who get an idea of what to look for by mulling over existing data, I began to get ideas from things I was seeing and hearing on the street" (Duneier 1999, 341 quoted in Pachirat 2018, 32). Accordingly, we contend that the critical practice of reflecting on the process of discovery need not be a negative one, as it can produce highly insightful understandings of why the answers we find in our research look the way they do (Tufekci 2014; Hargittai 2015). And, at times, this practice may provide insights that *should* question all of our intellectual communities' approaches. Regardless of what methodology we apply, if research on violent contexts exclusively focuses on research participants that are accessible with relative ease, said research may end up playing into the existing power dynamics (Bell-Martin and Marston 2019), studying vulnerable populations instead of those repressive forces to whom they are vulnerable to.

The Ethical Dilemmas of Absorbing Information: Where Participant Observation and Digital Information Scraping Intersect

Participant observation shares some of the ethical dilemmas of digital information "scraping." Like researchers working with large data sources, ethnographers record the traces of everyday life. For example, big data may retrace the footsteps of people through their GPS history, phone applications, or social media networks (Zook et al. 2017, 4). Ethnographers, on the other hand, may retrace footsteps (literally) through visual clues encountered with material objects and in spaces, by following or watching people's behaviors first hand, or by listening to stories and social interactions. Either way, participants leave these traces behind during the normal course of their lives, without the intent of providing researchers with a source of information or knowledge. While ethnographers do ask for and receive permission to be present in social settings and generally introduce their purposes openly, there is no formal consent process that recurs with every observation; this consent process is generally limited to interview settings. Eyes, ears, skin, and nose record details and ideas through the prism of our own humanity and experience but ubiquitously. Thus, the nature of participant observation complicates informed consent and right of withdrawal, similar to misgivings expressed by "big data" researchers about their own access (e.g., Zook et al. 2017). It also blurs the boundaries between private/public, requiring careful contextualization of observations to disentangle the ethics of research based upon them (Zook et al. 2017). As a tradition of

feminist scholarship has long acknowledged in fieldwork settings,[10] ultimately, the decisions about how these boundaries are maintained or transgressed are structured by power, whether we are dealing with information gleaned through new technology or old-fashioned participant observation (Crawford and Finn 2015, 499).

In response to these dilemmas, researchers working with large amounts of data have been racing to develop ethical norms to guide their work, and ethnographers would do well to notice and contemplate the applicability of these norms to their own field. Indeed, the uncertainty produced by rapid technological change is generating increasingly nuanced understandings of privacy and research consent. As discussed by Zook et al. (2017), privacy is not binary. In other words, the privacy preferences of individuals are situational, fluid, and culturally constructed (Crawford and Finn 2015, 498). Ethnographers have long argued that rich contextualized information about social setting and participants is necessary to make analytical judgments about the reliability of interpretivist research. However, quantitative researchers are highlighting how rich contextualized information about each data stream and the unique vulnerabilities of the individual participant is also necessary to make ethical judgments about transparency and harm of research, not just its analytical credibility (Zook et al. 2017). Quantitative researchers have begun to grapple with the ways that the very meaning of privacy evolve over time for society, as well as during the lifespan of individuals who may suffer unjust lingering impact from previous choices and exposures.

In addition to providing a parallel practice that ethnographers can look into as a mirror and learn from, "big data" also potentially puts qualitative materials gathered by ethnographers to new and unintended purposes. Given the rapid evolution of data extraction and aggregation techniques in which multiple sources of information can be cross-referenced, ethnographers may encounter new difficulties ensuring the de-identification of their field notes in the future. Even anonymized field notes might be re-identified as technology advances. Thus, informed consent in ethnographic projects, even under circumstances with formal process and ongoing dialogue with participants, can be rendered problematic by the encroachment of new technologies that improvise on qualitative research for new purposes, perhaps beyond the control of the ethnographer. The uncertainty produced by new technologies impacts the ethics of transparency for qualitative research. Ethnographers must assume that their information, including traces of their own lives and practices that had not been intended as public, such as their own movements tracked by a phone app during fieldwork, might be collected for aggregation alongside other data streams and thereby create problems for the de-identification of participants. In such a situation the mere presence of an ethnographer, even without any field notes, could expose participants to violent retaliation. Thus, ethnographers can no longer assume that they can maintain complete control over the privacy of all data generated by their presence in the field, and increasingly they must grapple with the idea that perhaps some research is too dangerous to conduct. It therefore behooves all ethnographers to stay abreast of debates and developments in "big data" research.

The answers to these dilemmas provided in the literature on "big data" do not invite complacency, but instead urge all researchers to continue to consider each case of transparency and information individually. For example, to navigate the context of uncertainty produced by rapid technological change, Zook et al. (2017, 7) advocate flexibility rather than strictly rule bound research behavior. Ethics for research that renders informed consent and public/private boundaries ambiguous and requires contextualized situational knowledge for judgment, such as that conducted with either quantitative or participant observation, cannot be easily bureaucratized

---

[10] See, for example, Golde 1970; Wolf 1992; Bell, Caplan, and Karim 1993; Miles and Crush 1993; Wolf 1996; Rose 1997; Warren and Hackney 2000; Mountz 2007; Gibson-Graham 2008; Hedge 2009; Berry et al. 2017.

by an institutional review board process with strict one-size-fits-all guidelines. Navigating uncertainty requires ongoing struggle, improvisation on rules, and reflexivity (Brigden, forthcoming).

Conclusion: Progress, Public Access, and Policy Relevance

In this article, we have argued the debate surrounding data access extends possibilities for building bridges between academic research communities at this critical historical juncture. These academic debates have become particularly heated, as they unfold within a political moment marred by public distrust and skepticism of traditional knowledge institutions and scientific expertise. Instead of focusing on areas of disagreement, we hope to demonstrate here that widely disparate methodological communities working on similar objects of interest—in this case: violence— can find important similarities in the ethical dilemmas they face and, furthermore, can learn from each other's advances in identifying pitfalls and setting standards in practices to avoid those pitfalls.

In this context, both ethnography and quantitative research require deeper reflection on the ethics of transparency for research on violence. Because of their shared global political conditions and the vulnerabilities of their participants, this deeper reflection necessitates dialogue across methodological positions. Rapid technological evolution and the urgency of humanitarian crises is forcing "big data" researchers to recognize the need for flexible ethics that respond to dynamic and specific real-world settings. Engaging with the standards and discussions interpretivists have been holding for many decades will help quantitative researchers navigate many of these new and urgent ethical dilemmas. At the same time, their navigation of intense uncertainty and urgency provides lessons for ethnographers as well. In an important sense, "big data" provides a mirror for ethnographic practices, such as participant observation, that blurs boundaries between public/private. The rapid evolution of technology has spurred new thinking about these boundaries that may prove instructive for ethnographers to engage. Finally, because their own outputs could unwittingly be utilized by researchers with these powerful aggregation tools, ethnographers have a duty to monitor debates over the ethics of big data, as well as debates within their own qualitative field.

The dialogue presented here represents a starting place, highlighting points of compatibility between our thinking about the ethics of studying violence. As we conclude this piece, we continue to sit uncomfortably with the potential harms caused by our research, the questions we ask, the boundaries we set, and our own transparency choices, both in the immediate moment and looking into an unknown future. Indeed, for a long time, this has been a concern for us. In a recent online symposium published by *International Studies Quarterly*, Brigden (2017) explains these concerns: "In the sort of nights that keep researchers awake, puzzling over the politics of their own role in the theater of borders, I have often felt wary about the relative ease with which I could interview a vulnerable, clandestine population, migrants whose very survival depends on being undocumented and illegible to the State. It is much more difficult to witness the inner workings of the State itself; the entry points to the State are more closely guarded than the entry points to the underground. Power and politics make such access and the task of 'ethnography of the state' inherently difficult, but ever more important (Mountz 2007, 10)."

Thus, Brigden (2017) felt uneasy with the fact that, in the context of contemporary border policing, it is easier to access a smuggling passage than to witness the secret inner workings of the state. The state, like smugglers, functions in the shadows, creating harm to vulnerable people. That harm is *clandestine*, defined by Coutin (2005) as hidden from view but known. Research that makes the state legible to everyday people, rather than everyday people legible to the state, can be more difficult to undertake, and that is no coincidence; it is a product of power

(Mountz 2007). Indeed, this asymmetry is a major limitation of ethical ethnographic methods, especially in areas of national security concern where the state guards itself so tightly (Brigden and Mainwaring 2019). As it turns out, it is also a major limitation and ethical problem for "big data" methods, where access is also mediated by power, such as through state restrictions or technology companies.

In both of our work, we have examined resistance to state violence by vulnerable people. These are situations in which state policy is explicitly designed to produce harm, rather than mediate it. In Brigden's area of work, governments attempt to deter migration by increasing the criminal penalties and lethal risks to border crossers, including asylum seekers. Governments also attempt to criminalize and suppress humanitarian work, including information gathering about violations of migrant and refugee rights, with these border crossers. In Gohdes's work, governments target, threaten, and murder human rights advocates, journalists, and political rivals; erase evidence; and criminalize actors (including displaced people) who wish to report or monitor these misdeeds.

From institutional affiliations on opposite sides of the Atlantic, we have watched with horror renewed nationalism, xenophobia, and authoritarianism in both the Americas and Europe. The transnational march of these social and political forces has intensified the urgency with which we question the very purpose of our research, and it has raised questions about our relationship to our own governments, which may seek to either infiltrate or repress our research. In the aftermath of the war on terror and the new age of surveillance it has ushered in, the deepening of violence against political refugees and other displaced populations, in particular, has collapsed the boundaries between our "field" and our home.

From this vantage point, we conclude by contemplating how a thick transparency might better serve our larger normative commitments. In so doing, we urge political scientists to jettison calls for "policy relevance" (narrowly understood) in favor of "people's relevance" or perhaps even "people's resistance." When policies are designed to hurt, the last thing that we want is for our research to be "policy relevant." A broader, richer understanding of transparency, informed by an understanding of the inherently political, unpredictable, and socially constructed nature of privacy and vulnerability, better serves this broader agenda of people's relevance, regardless of the methodological commitments of the researcher. With this ethics in mind, we invite ethnographers and quantitative researchers to talk together about transparency.

Acknowledgements

References

AAA Executive Board. 2007. Statement on the Human Terrain System Project. American Anthropological Association. October 31, 2007. http://www.aaanet.org/issues/policy-advocacy/Statement-on-HTS.cfm.

AUTESSERRE, SEVERINE. 2010. *The Trouble with the Congo: Local Violence and the Failure of International Peacebuilding*. Cambridge: Cambridge University Press.

BELL, DIANE, PAT CAPLAN, AND WAZIR JAHAN KARIM. 1993. *Gendered Fields: Women, Men and Ethnography.* London: Routledge Press.

BELL-MARTIN, REBECCA V., AND JEROME F. MARSON, JR. 2019. "Confronting Selection Bias: The Normative and Empirical Risks of Data Collection in Violent Contexts." *Geopolitics.* https://doi.org/10.1080/14650045.2019.1659780.

BERRY, MAY J., CLAUDIA CHÁVEZ ARGÜELLES, SHANYA CORDIS, SARAH IHMOUD, AND ELIZABETH VELÁSQUEZ ESTRADA. 2017. "Toward a Fugitive Anthropology: Gender, Race, and Violence in the Field." *Cultural Anthropology* 32 (4): 537–65.

BLEICH, ERIK, AND ROBERT J. PEKKANEN. 2015. "Data Access, Research Transparency, and Interviews: The Interview Methods Appendix." *Qualitative and Multi-Method Research: Newsletter of the American Political Science Association's QMMR Section* 13 (1): 8–13.

BOURGOIS, PHILIPPE. 2001. "The Power of Violence in War and Peace: Post–Cold War Lessons from El Salvador." *Ethnography* 2 (1): 5–34.

BOYD, DANAH, AND KATE CRAWFORD. 2012. "Critical Questions for Big Data: Provocations for a Cultural, Technological and Scholarly Phenomenon." *Information, Communication & Society* 15 (2): 662–79.

BREEN-SMYTH, MARIE. 2019. "Interviewing Combatants: Lessons from the Boston College Case." *Contemporary Social Science*: 1–17. https://doi.org/10.1080/21582041.2019.1637533.

BRIGDEN, NOELLE K. Forthcoming. "Methods that Mirror Migration: Ethics and Entanglement En Route." In *Research Methods in Critical Security Studies, 2nd ed.*, edited by Mark B. Salter, Can E. Mutlu, and Philippe M. Frowd. New York: Routledge Press.

———. 2017. "On Metaphors, Methods and Motion: A Response." Symposium: Walking with Migrants: Ethnography as Method in International Relations. *International Studies Quarterly* Harvard Dataverse, V1 2017. https://doi.org/10.7910/DVN/WUK8AN.

———. 2018a. *The Migrant Passage: Clandestine Journeys from Central America.* Ithaca, NY: Cornell University Press.

———. 2018b. "A Visible Geography of Invisible Journeys: Information, Representation and the Politics of Survival along the Migrant Trail from Central America." *International Journal of Migration and Border Studies* 4 (1/2): 71–88.

BRIGDEN, NOELLE K., AND MIRANDA HALLETT. 2020. "Fieldwork as Social Transformation: Place, Time, and Power in a Violent Moment" *Geopolitics.* https://doi.org/10.1080/14650045.2020.1717068.

BRIGDEN, NOELLE K., AND CETTA MAINWARING. 2019. "Ethnography and Global Politics: Engaging International Relations from Below." Working draft, Marquette University.

BÜTHE, TIM, AND ALAN M. JACOBS. 2015. "Transparency in Qualitative and Multi-method Research: Introduction to the Symposium." *Qualitative & Multi-Method Research* 13 (1): 2–8.

CEDERMAN, LARS-ERIK, AND KRISTIAN SKREDE GLEDITSCH. 2009. "Introduction to Special Issue on 'Disaggregating Civil War.'" *Journal of Conflict Resolution* 53 (4): 487–95.

CRAMER, KATHERINE. 2015. "Transparent Explanations, Yes. Public Transcripts and Fieldnotes, No: Ethnographic Research on Public Opinion." *Qualitative & Multi-Method Research* 13 (1): 17–20.

CRAWFORD, KATE, AND MEGAN FINN. 2015. "The Limits of Crisis Data: Analytical and Ethical Challenges of Using Social and Mobile Data to Understand Disasters." *GeoJournal* 80 (4): 491–502.

CRIADO-PEREZ, CAROLINE. 2019. *Invisible Women: Exposing Data Bias in a World Designed for Men.* New York: Abrams Press.

COUTIN, SUSAN BIBLER. 2005. "Being en Route." *American Anthropologist* 107 (2): 195–206.

DINUR, IRIT, AND KOBBI NISSIM. 2003. "Revealing Information while Preserving Privacy." In Proceedings of the Twenty-Second ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems. Association for Computing Machinery, 202–10.

DUNEIER, MITCHELL. 1999. *Sidewalk.* New York: Farrar, Straus and Giroux.

DURÁN-MARTÍNEZ, ANGÉLICA. 2018. *The Politics of Drug Violence: Criminals, Cops and Politicians in Colombia and Mexico.* New York: Oxford University Press.

DWORK, CYNTHIA. 2011. "Differential Privacy." In *Encyclopedia of Cryptography and Security*, edited by Henk C. A. van Tilborg and Sushil Jajodia, 338–40. New York: Springer.

ECK, KRISTINE, AND LISA HULTMAN. 2007. "One-Sided Violence against Civilians in War: Insights from New Fatality Data." *Journal of Peace Research* 44 (2): 233–46.

EMERSON, ROBERT M., RACHEL I. FRETZ, AND LINDA L. SHAW. 1995. *Writing Ethnographic Fieldnotes.* Chicago: University of Chicago Press.

EUROPEAN COMMISSION. 2019. Horizon 2020 Online Manual: Ethics. Accessed October 30, 2019. https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/ethics_en.htm.

FEARON, JAMES D., AND DAVID D. LAITIN. 2003. "Ethnicity, Insurgency, and Civil War." *American Political Science Review* 97 (1): 75–90.

Fuji, Lee Ann. 2009. *Killing Neighbors: Webs of Violence in Rwanda.* Ithaca, NY: Cornell University Press.

Gibson-Graham, J. K. 2008. "Diverse Economies: Performative Practices for 'Other Worlds.'" *Progress in Human Geography* 32 (5): 613–32.

Glaser, April. "Palantir Said It Had Nothing to Do with ICE Deportations: New Documents Seem to Tell a Different Story." Slate. May 2, 2019. https://slate.com/technology/2019/05/documents-reveal-palantir-software-is-used-for-ice-deportations.html.

Goffman, Erving. 1959. *The Presentation of Self in Everyday Life.* Garden City, NY: Doubleday.

Gohdes, Anita R. Forthcoming. "Repression Technology: Internet Accessibility and State Violence." *American Journal of Political Science.* https://doi.org/10.1111/ajps.12509.

Golde, Peggy. 1970. "Introduction." In *Women in the Field: Anthropological Experiences*, edited by Peggy Golde, 1–18. Berkeley: University of California Press.

Gough, Kathleen. 1968. "Anthropology and Imperialism." *Economic and Political Weekly. Monthly Review* 25 (31): 12–27.

Harff, Barbara. 2003. "No Lessons Learned from the Holocaust? Assessing Risks of Genocide and Political Mass Murder since 1955." *American Political Science Review* 97 (1): 57–73.

Hargittai, Eszter. 2015. "Is Bigger Always Better? Potential Biases of Big Data Derived from Social Network Sites." *The Annals of the American Academy of Political and Social Science* 659 (1): 63–76.

Hedge, R. S. 2009. "Fragments and Interruptions: Sensory Regimes of Violence and the Limits of Feminist Ethnography." *Qualitative Inquiry* 15 (2): 276–96.

Idler, Annette. 2019. *Borderland Battles: Violence, Crime and Governance at the Edges of Colombia's War.* Oxford: Oxford University Press.

Krystalli, Roxani. 2018. "Negotiating Data Management with the National Science Foundation: Transparency and Ethics in Research Relationships." Memo. Accessed June 28, 2019. http://interpretationandmethod.com/wp-content/uploads/2015/10/Krystalli-NSF-Data-Sharing-Memo_ForPosting_March2019.pdf.

Leeds, Brett Ashley. 2003. "Do Alliances Deter Aggression? The Influence of Military Alliances on the Initiation of Militarized Interstate Disputes." *American Journal of Political Science* 47 (3): 427–39.

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

Lyall, Jason, Graeme Blair, and Kosuke Imai. 2013. "Explaining Support for Combatants During Wartime: A Survey Experiment in Afghanistan." *American Political Science Review* 107 (4): 679–705.

Mitts, Tamar. 2019. "From Isolation to Radicalization: Anti-Muslim Hostility and Support for ISIS in the West." *American Political Science Review* 113 (1): 173–94.

Miles, Miranda, and Jonathan Crush. 1993. "Personal Narratives as Interactive Texts: Collecting and Interpreting Migrant Life-Histories." *The Professional Geographer* 45 (1): 84–94.

Mountz, Alison. 2007. "Smoke and Mirrors: An Ethnography of the State." In *Politics and Practice in Economic Geography*, edited by Adam Tickell, Eric Sheppard, Jamie Peck and Trevor Barnes, 38–48. Thousand Oaks, CA: Sage.

Murgia, Madhumita. 2019. "Who's Using Your Face? The Ugly Truth about Facial Recognition." *Financial Times*, September 18, 2019. https://www.ft.com/content/cf19b956-60a2-11e9-b285-3acd5d43599e.

Narayanan, Arvind, and Vitaly Shmatikov. 2009. " De-anonymizing Social Networks." 30th IEEE Symposium on Security and Privacy, Berkeley, CA, 2009, 173–87. https://ieeexplore.ieee.org/abstract/document/5207644.

Oneal, J. R., and B. Russett. 2001. *Triangulating Peace: Democracy, Interdependence, and International Organizations.* New York: Norton.

Pachirat, Timothy. 2013. *Every Twelve Seconds: Industrialized Slaughter and the Politics of Sight.* New Haven, CT: Yale University Press.

———. 2018. *Among Wolves: Ethnography and the Immersive Study of Power.* New York: Routledge Press

Parkinson, Sarah Elizabeth. 2013. "Organizing Rebellion: Rethinking High-Risk Mobilization and Social Networks in War." *American Political Science Review* 107 (3): 418–32.

Parkinson, Sarah Elizabeth, and Elisabeth Jean Wood. 2015. "Transparency in Intensive Research on Violence: Ethical Dilemmas and Unforeseen Consequences." *Qualitative & Multi-Method Research* 13 (1): 22–27.

Preston, Alex. 2014. "The Death of Privacy." *Guardian*, August 3, 2014. https://www.theguardian.com/world/2014/aug/03/internet-death-privacy-google-facebook-alex-preston.

Price, Megan, and Patrick Ball. 2014. "Big Data, Selection Bias, and the Statistical Patterns of Mortality in Conflict." *SAIS Review of International Affairs* 34 (1): 9–20.

Price, Megan, Anita Gohdes, and Patrick Ball. 2014. "Updated Statistical Analysis of Documentation of Killings in the Syrian Arab Republic." *Human Rights Data Analysis Group.* Accessed June 28, 2019. https://ohchr.org/Documents/Countries/SY/HRDAGUpdatedReportAug2014.pdf.

———. 2016. "Technical Memo for Amnesty International Report on Deaths in Deten-
tion." *Human Rights Data Analysis Group.* Accessed June 28, 2019, https://www.hrdag.org/wp-
content/uploads/2016/08/HRDAG-AI-memo-6.pdf.

RESPONSIBLE DATA 2019. "What Is Responsible Data?" Accessed June 28, 2019. https://
responsibledata.io/what-is-responsible-data/.

ROSE, GIDEON. 1997. "Situating Knowledges: Positionality, Reflexivities and Other Tactics." *Progress in
Human Geography* 21 (3): 305–20.

SCHATZ, EDWARD. 2009. *Political Ethnography: What Immersion Contributes to the Study of Power,* edited by Ed-
ward Schatz. Chicago: University of Chicago Press.

SHAPIRO, JACOB N., AND NILS B. WEIDMANN. 2015. "Is the Phone Mightier Than the Sword? Cellphones and
Insurgent Violence in Iraq." *International Organization* 69 (2): 247–74.

SOLOVE, DANIEL J. 2012. "Introduction: Privacy Self-Management and the Consent Dilemma." *Harvard
Law Review* 126: 1880–903.

THALER, KAI. 2019. "Reflexivity and Temporality in Researching Violent Settings: Problems
with the Replicability and Transparency Regime." *Geopolitics* 1–27. https://doi.org/10.1080/
14650045.2019.1643721.

TUFEKCI, ZEYNEP. 2014. "Big Questions for Social Media Big Data: Representativeness, Validity and Other
Methodological Pitfalls." *Proceedings of the Eighth International AAAI Conference on Weblogs and Social
Media.*

WARREN, CAROL A. B., AND JENNIFER KAY HACKNEY. 2000. *Gender Issues in Ethnography,* 2nd ed., Qualitative
Research Methods Series 9. Thousand Oaks, CA: Sage Publications.

WEDEEN, LISA. 2008. *Peripheral Visions: Publics, Power and Performance in Yemen.* Chicago: University of
Chicago Press.

WEDEEN, LISA. 1999. *Ambiguities of Domination: Politics, Rhetoric and Symbols in Contemporary Syria.* Chicago:
University of Chicago Press.

WEIDMANN, NILS B. 2013. "The Higher the Better? The Limits of Analytical Resolution in Conflict Event
Datasets." *Cooperation and Conflict* 48 (4): 567–76.

WOLF, DIANE L. 1996. *Feminist Dilemmas in Fieldwork,* edited by Diane L. Wolf. Boulder, CO: Westview Press.

WOLF, MARGERY. 1992. *A Thrice Told Tale: Feminism, Postmodernism & Ethnographic Responsibility.* Stanford,
CA: Stanford University Press.

WOOD, ELISABETH JEAN. 2003. *Insurgent Collective Action and Civil War in El Salvador.* Cambridge: Cambridge
University Press.

ZOOK, MATTHEW, SOLON BAROCAS, DANAH BOYD, KATE CRAWFORD, EMILY KELLER, SEETA PEÑA GANGADHARAN, AND
ALYSSA GOODMAN et al. 2017. "Ten Simple Rules for Responsible Big Data Research." *PLOS Computa-
tional Biology* 13 (3): 1–10.