

## Thinking outside the (archive) box: discovering data in the field

The previous two chapters explored the broad challenges researchers face when preparing for fieldwork and operating in the field. This chapter and the next four focus more specifically on data-collection techniques that political scientists use while conducting field research. As we have emphasized, fieldwork entails much more than collecting information. Nonetheless, data generation is unarguably a central component of field research.<sup>1</sup> Our objective in Chapters 5 through 9 is to contextualize these data-collection techniques, considering what challenges their deployment in the field entails and offering strategies to address them. More broadly, the chapters discuss the techniques' advantages and disadvantages, consider how they can be combined, and address how to evaluate the evidentiary value of the data they generate. Most critically, the chapters demonstrate how collecting data using each technique contributes to the accumulation of knowledge and development of theory in political science.

We begin with perhaps the least well-specified data-collection technique employed by political scientists who conduct field research: gathering pre-existing materials. This mode of collecting information is fundamentally different from the more interactive techniques covered in the next four chapters – interviews, site-intensive methods, surveys, and experiments. Of course, collecting pre-existing materials often involves *some* interaction with people – in order to gain access to them (speaking with an archivist, asking to peruse a private collection, or asking permission to take a photograph), to sample them, or to capture them. As with all interactions in the field, scholars should critically reflect on these more fleeting exchanges and their implications for the information being collected. The basic difference between searching for, identifying, discovering,

<sup>1</sup> Our definition of “data,” introduced in Chapter 1, bears rearticulating here: for us, data are materials (information or observations) that have been collected and processed by a researcher – considered in context and assigned some analytic significance – such that they can be employed in his analysis.

and gathering pre-existing materials in the field and the other modes of data collection we discuss in this book is that gathering pre-existing materials does not entail *generating* sources or information in the way that conducting interviews or experiments does.

Some might equate collecting pre-existing materials with archival research (e.g., gathering government documents, other historical matter, or newspaper articles from a defined location). However, as our chapter title suggests, pre-existing materials of use to a scholar's project may reside in many locations beyond archives. Indeed, a central contribution of this chapter is to encourage scholars to think creatively about *other* data sources they can exploit in the field. Statistical datasets, maps, NGO advocacy papers, political party platforms, photos, posters, pamphlets, brochures – even bumper stickers and graffiti (to name just a few), can all be important sources of data that can help researchers build their arguments. Some of these diverse data sources can be found in specialized locations, and their discovery can be purposive and directed; others surround researchers in the world they experience every day, and their discovery is often more accidental or serendipitous.

Despite this potential richness, if collecting pre-existing materials is mentioned at all when fieldwork is taught in graduate methods courses, short courses at the APSA annual meeting, or the IQMR, generally only archival research is discussed.<sup>2</sup> This gap may result from the generalized impression that pre-existing materials found outside of archives cannot yield usable data, or that the process of collecting them is either straightforward (and thus does not need to be taught) or chaotic (and thus cannot be taught). The lacuna may also spring from a view that few political scientists carry out research projects based *exclusively* on pre-existing materials. Alternatively, the lack of attention to collecting such materials may be due to a belief that appropriate archival methods differ so much across research questions and field sites that none is broadly applicable – or because the literature addressing archival methodology (even in the discipline of history) is somewhat limited.<sup>3</sup>

Regardless of its origins, the gap is unfortunate given the important role such materials often play in political science inquiry. Most scholars collect

<sup>2</sup> Dedicated workshops on archival research, such as the annual Summer Institute on Conducting Archival Research (SICAR) associated with the Program on Conducting Archival Research (POCAR) at George Washington University, also exist.

<sup>3</sup> Personal communication with three senior historians, March 14, 2011; July 23, 2012; July 24, 2012. The consensus among these scholars was that, while archival research is often discussed, and history graduate students explore texts grounded in archival research through their coursework, ultimately they learn archival work "by encounter."

pre-existing materials at some point in their fieldwork. The results of our survey of US political science faculty bear out this assertion: 70 percent of field research projects in our sample drew on various sorts of media sources when collecting data, 68 percent involved the collection of books and articles, 64 percent included archival research. In terms of numerical sources, 49 percent entailed compiling quantitative data through means other than surveys and experiments,<sup>4</sup> and 32 percent involved collecting existing datasets. These sorts of pre-existing materials, which political scientists so often instinctively pick up, can significantly shape how they think about their question and their topic. Depending on the project, not being more systematic about collecting and considering such materials, and more reflective about how they influence our thinking, can mean lost opportunities. Indeed, perhaps due to a generalized impression that many of these types of sources (perhaps with the exception of archival documents) do not constitute “real data,” they are infrequently cited in analyses, and the role they play in knowledge generation is rarely acknowledged in scholars’ publications. This omission represents a threat to research transparency, leaving the empirical base of researchers’ analyses underspecified.

Our goal in this chapter, then, is to offer a provocative primer on identifying and collecting pre-existing materials, with a particular emphasis on archival research. The chapter proceeds as follows. In the next section, we encourage scholars to think creatively about the many sources and types of pre-existing data that may be useful for their work, arguing that “thinking outside the (archive) box” can pay big analytic dividends. We also consider the differences between collecting pre-existing materials and engaging in the more interactive data-collection techniques discussed later in the book, and show how the two modes of data collection can be productively combined. We then demonstrate how pre-existing materials collected in the field can contribute to theory building in political science. Next we consider some of the challenges inherent in collecting pre-existing materials – identifying and sampling sources, addressing missing information, and evaluating and identifying bias – and suggest some strategies to address them; this section also raises some ethical considerations. The chapter’s last substantive section briefly discusses preparing for and engaging in archival research, highlighting some unique aspects of archival settings.

<sup>4</sup> In some of these projects, scholars may have been building collections of new data rather than collecting pre-existing data.

## Collecting pre-existing materials, and comparing and combining their collection with more interactive data collection

Pre-existing sources of innumerable types can be found in practically any field site. Indeed, no matter what the substantive or temporal focus of their projects, most researchers will at some point gather artifacts of popular culture, government documents, maps, public opinion data, political party brochures, newspaper articles, or other types of pre-existing materials. Table 5.1 outlines a suggestive range of such sources, noting some of the locations in which they might be found.<sup>5</sup>

As the table suggests, pre-existing materials come in many forms. We hesitate to make much of the conventional distinction between primary and secondary sources, as many pre-existing materials might qualify as both.<sup>6</sup> A report by an NGO might contain both raw facts of a political issue and analysis of it, for example. Distinctions can sometimes be drawn, however, with regard to how “official” or “formal” pre-existing materials are. For instance, a statistical dataset acquired from a government ministry may be more official, a garment reprinted with a political leader’s photograph happened upon in a bazaar may be less official, and a set of position papers found scattered about various NGOs may be somewhere in between. Nonetheless, all can be useful to political scientists.

Pre-existing data sources can also vary greatly in terms of accessibility. Sometimes restrictions are in place, and other times protective gatekeepers (political actors or researchers) stand between data and researchers. Moreover, pre-existing materials can reflect particular agendas, contain varied biases, or intentionally emphasize (and deemphasize) particular aspects or dynamics of the phenomenon of interest. All of this can affect their deployment as evidence in scholars’ analyses. For example, public records may stress administrative considerations and downplay political pressures on officials, while the press and officials’ correspondence or memories may do the opposite (Tosh 2000, 59).

Broadly speaking, scholars may identify pre-existing materials of potential relevance to their work in two sets of ways. On the one hand, sometimes

<sup>5</sup> Tosh (2000, 27–47) outlines and discusses, from a historian’s viewpoint, the main categories of documentary material.

<sup>6</sup> By “primary” sources, we mean “raw” or “original” sources contemporary to the event or dynamic to which they relate; by “secondary” sources, we mean written interpretations of the past – an analytical product that may draw on primary sources (Thies 2002, 356). Tosh (2000, 29–31) discusses the blurry line, and relationship, between the two, and emphasizes historians’ preference for primary sources.

## Collecting pre-existing materials

**Table 5.1 Examples of pre-existing source materials and where to find them**

| Source materials   | Locations   |
|--|---|
| <ul style="list-style-type: none"> <li>● Posters, postcards, bumper stickers, graffiti, invitations, flyers, popular-culture visual or audio materials (printed cloth, art, songs, etc.), and other informal/unofficial materials</li> </ul>   | The broad research setting  |
| <ul style="list-style-type: none"> <li>● Maps, satellite imagery, geographic data</li> <li>● Official documents, files, reports (diplomatic, public policy, propaganda, etc.)</li> <li>● Government statistics (e.g., on population, sewage systems, roads, GDP, budgets, immigration, trade, poverty)</li> <li>● Correspondence, memoranda, communiqués, queries, complaints</li> <li>● Parliamentary proceedings</li> <li>● Minutes from intra-governmental and inter-governmental meetings</li> <li>● Testimony in public hearings</li> <li>● Speeches, press conferences</li> <li>● Military records</li> <li>● Court records</li> <li>● Legal documents (charts, wills, contracts)</li> <li>● Chronicles, autobiographies, memoirs, travel logs, diaries</li> </ul> | Archives (national or local, government or private)<br>Government agencies/entities (e.g., congressional offices, ministries, etc., at various levels)        |
| <ul style="list-style-type: none"> <li>● Private papers</li> <li>● Brochures, posters, flyers</li> <li>● Press releases, newsletters, annual reports</li> <li>● Records, papers, directories</li> <li>● Internal memos, reports, meeting minutes</li> <li>● Position or advocacy papers, mission statements</li> <li>● Party platforms</li> </ul>  | Individuals or organizations (e.g., political parties, unions, businesses, hospitals, schools, religious entities, interest groups, universities, NGOs, etc.) |
| <ul style="list-style-type: none"> <li>● Radio broadcasts (transcripts)</li> <li>● TV programs (transcripts)</li> <li>● Magazines, newspapers</li> <li>● Electronic media</li> </ul>   | Media   |
| <ul style="list-style-type: none"> <li>● Published collections of documents, gazetteers, yearbooks, etc.</li> <li>● Private papers</li> <li>● Books, articles, dissertations, working papers</li> </ul>  | Libraries (university, national, local)   |

the fieldwork setting serendipitously *produces* such materials, or offers cues and clues that point to sources scholars might not have anticipated gathering. For instance, a scholar might be handed a flyer at a demonstration, or glimpse a telling bumper sticker on a passing car or revealing graffiti on a wall. One researcher studying informal workers in Peru befriended a street vendor who began to lend her documents from a large archive he had created over the previous decade on the legal and social evolution of street vending in his district of Lima; ultimately, she photocopied 850 pages that served as a “data goldmine” for her project.<sup>7</sup> Here the identification and collection of pre-existing materials are by nature unsystematic, unplanned, and accidental; scholars are generally not following pre-determined rules about what to collect (and often have little choice).<sup>8</sup> In order to benefit from these opportunities scholars need do little more than remain alert and engaged, actively monitoring the world around them – and, of course, reflect critically on why they may have happened upon, or been provided, certain materials.

On the other hand, pre-existing materials can be collected more purposefully. Scholars often set their sights on certain types of sources – e.g., archival documents, court records, legislative minutes, newspaper articles, datasets, press releases, or secondary sources (historical accounts, for instance) – and go about acquiring them in a much more directed way.<sup>9</sup> They might do so by requesting them from the entity that produced or stores them; downloading them from the relevant organization’s web site; or obtaining them from a library or electronic academic journal. For instance, one comparativist captured the entire population of television advertisements run by the main presidential candidates in several recent elections in three countries, and also acquired videos containing every television spot from several previous presidential elections in each country.<sup>10</sup> More specifically, archives, newspapers, and other types of sources themselves can be searched and selected systematically.

Even in view of this heterogeneity, the processes of collecting pre-existing materials (and the materials themselves) have certain similarities that set them apart from more interactive data-collection techniques (and the

<sup>7</sup> Interview, DK-1, July 20, 2012.

<sup>8</sup> To be clear, we are not suggesting that these events are fully random; using flyers as an example, their distributors may choose to give them to some people and not to others on the basis of factors over which the researcher has little control, and of which he may often remain ignorant.

<sup>9</sup> The distinction is not iron-clad; scholars may happen upon particularly useful documents when they are searching for others, for instance.

<sup>10</sup> Interview, DK-3, July 27, 2012.

information gathered through employing them). Pre-existing materials tend to be inanimate objects (maps, newspaper articles, documents, statistics). Accordingly, as noted earlier, while identifying, selecting, and capturing pre-existing sources may involve human interaction, extracting information and observations from them does not. Relatedly, while people or groups can be affected in many ways by a scholar interacting with them to obtain information, pre-existing materials cannot: documents do not react to being gathered and read in the way that interview respondents react to being prodded for answers to questions.<sup>11</sup> At a minimum then, the bias inherent in a pre-existing source notwithstanding, it becomes no less valid through the collection process.

These sources' pre-existing nature and non-reactivity can have downsides, however. A researcher cannot ask a dataset to produce another column containing the type of information he was hoping to find, or ask a piece of graffiti to tell him more. There are more potential silences when extracting information from pre-existing materials, and it may be more difficult for the researcher to understand sources with which he cannot interact.<sup>12</sup> Relatedly, such sources' pre-existing nature means that they are remnants and artifacts that reflect the choices and interests of *other* scholars or actors – influences that often remain unknown to a researcher who pulls them from an archive box decades or centuries later.<sup>13</sup> Further, many long days, weeks, and months in an archive may fail to yield the precise information a scholar needs and could access if he could just *talk* to the relevant actors. There is also more likely to be an intermediary between the researcher and pre-existing materials – a clerk in an archive, the individual who assembled a relevant dataset, or the person who took the meeting minutes of interest – lengthening the distance between him and the actual information source. All of these properties of pre-existing sources must be taken into account as scholars seek to assess their evidentiary value and deploy them to underpin their analyses.

As technology advances, more pre-existing materials are continually becoming available and scholars' relationships with their sources are evolving. The web can be thought of as an electronic archive – or an archive

<sup>11</sup> Of course, scholars *do* react to data sources, and may even see the same data in a different light over time.

<sup>12</sup> See Hill (1993, 68) for an evocative description of the difference between interactive and archival research; interview, LM-7, September 20, 2012.

<sup>13</sup> The coding of documents and interpretation of artifacts, etc., are active processes with plenty of latitude for choice.

of archives (for more on this notion, see Sentilles 2005). Online worlds, Facebook, YouTube, blogs, and other types of social media can be rich sources of information relevant to political science analyses.<sup>14</sup> And pre-existing data increasingly come paired with locational coordinates or other geographic information. The future will likely see more applications of these and other techniques for collecting information and generating data that do not centrally involve interaction with people. These techniques, of course, test the boundaries of our definition of fieldwork, the central component of which is “being there.” Optimizing understanding, we would argue, requires developing creative ways to combine such digital materials with on-the-ground investigation or other means of engaging with the contexts from which they hail.

More generally, combining pre-existing materials with information collected using more interactive techniques – which can help scholars to engage in triangulation, one of our six principles of good field research – allows the strengths of the data gathered using one mode of collection to offset the weaknesses of those gathered using another, yielding rich intellectual payoffs. For instance, there is great potential for productive overlap and synergies between the site-intensive methods (SIM) discussed in Chapter 7 and the techniques discussed here. That chapter adopts an encompassing view of observation, discussing and encouraging both the sustained, focused kind of observation involved in ethnography and participant observation, *and* the less focused, less planned, and less sustained observation of context in which political scientists engage constantly while in the field. The observations about individuals and interactions made by scholars engaging in SIM, sometimes fleeting and always lacking in physical permanence, offer critical context for the concrete pre-existing materials scholars collect (often simultaneously), helping them to understand and interpret those sources.

Engaging in interviewing in tandem with collecting pre-existing materials can also be extremely productive. For scholars carrying out research on recent political events, the names of potential interview subjects – political figures, journalists who were present at important events, secretaries, note-takers, translators – may appear in newspaper or magazine articles. Other pre-existing sources, such as memoirs and archival documents, can also help scholars to identify people or groups whose involvement in the processes under study remained invisible, and whose voices had been absent from

<sup>14</sup> Chapter 4 offers examples, and Chapter 11 discusses how new technologies are shaping the future of field research.

conventional political narratives. Collecting and analyzing pre-existing materials can also help scholars prepare, and better capitalize on opportunities, for in-depth interviewing. For instance, the researcher who collected videos of presidential candidates' advertisement spots also conducted interviews with those candidates and others about their campaign strategies, and his familiarity with the videos allowed him to formulate more specific and better questions.<sup>15</sup> Another scholar with whom we spoke interviewed civic association leaders in South Africa about a membership recruitment flyer she had obtained at one of the association's other branches. The respondents' lack of awareness of the flyer revealed weaknesses in the association's overall organizational structure as well as political cross-pressures between the association and the ruling political party.<sup>16</sup>

Likewise, interview respondents can point researchers to archival sources they had not previously identified (or reveal that they have collections themselves!), or raise relevant questions that scholars can subsequently investigate through archival sources. For instance, in carrying out his project on city politics in Taipei, Read delved into almanacs of election records housed in Taiwan's national library to find supplementary information and verify facts that shaped or conditioned his findings from interviews with people involved in those elections. For another scholar who was coding newspapers for "taboo content" and identifying articles concerning political activism, interviews helped her to refine her coding strategy and categories, get a better sense of the history of and motivations for activism, and more fully understand the relationship between journalists and activists.<sup>17</sup> Information drawn from pre-existing materials can also help researchers better understand, cross-check, clarify, or fill in gaps in what they previously learned through interviews.<sup>18</sup> Finally and most specifically, interviewing archive gate-keepers and archivists can sometimes facilitate archival access by defusing tensions (drawing in those who could create obstacles for one's project) – as well as providing additional insight on the documents being gathered. Interviewing such individuals can also help scholars to understand where the organization whose archives they are accessing fits in the political scene, to fill in narrative holes and make sense of documents, and to test out hunches developed on the basis of those documents.

<sup>15</sup> Interview, DK-3, July 27, 2012.

<sup>17</sup> Interview, DK-5, July 31, 2012.

<sup>16</sup> Interview, LM-25, October 21, 2013.

<sup>18</sup> Interview, DK-1, July 31, 2012.

Pre-existing materials, then, come in many shapes and sizes, can be collected more or less purposively, and share several qualities that distinguish them from data collected via more interactive techniques. It can often be quite productive to combine the collection of pre-existing sources with more interactive modes of data gathering as such materials – particularly those happened upon serendipitously – often represent fragments of information whose analytic importance is only revealed through their combination with other observations. Even when combining is not possible (as occurs, for instance, when studies focus on dynamics that happened so far in the past that oral narratives about the event cannot be produced), scholars can still cross-check information by triangulating among different types of pre-existing materials.

---

### Using pre-existing materials to build theory

---

Despite the broad citing of archival sources in political science scholarship, some researchers may be skeptical about the evidentiary value of the information that can be drawn from pre-existing sources and the contribution it can make to political science inquiry. Some may even hold that the kinds of sources on which this chapter focuses (and archival sources in particular) *inhibit* broad theory building because the minutiae and descriptive detail they contain can lead scholars to tell a story that is either too partial or too nuanced. Our own experiences, and our empirical study of field research in the discipline and scholarship produced on the basis of fieldwork, suggest otherwise. Drawing on pre-existing materials of the types discussed here can add empirical depth to a research project and demonstrate strong knowledge of on-the-ground dynamics. Moreover, such materials can make a significant contribution to every stage of research – formulating research questions, selecting cases, conceptualizing and measuring key variables, generating potential explanations, and illuminating causal processes and mechanisms – and, thus, to theory building.

---

### Developing research questions

Information drawn from pre-existing materials concerning the details or dynamics of political events may help scholars to identify, refine, or more precisely formulate a research question; to confirm that a chosen question has merit (or does not); or to discover an entirely new question, problem, or

issue to be studied. For example, seeing a Communist party poster on a wall alerted a scholar of international relations in South Asia to how big an issue corruption in government was at the time of his research, and led him to focus on the issue in a sustained way. Doing so enriched his work considerably and enhanced its relevance, as corruption ultimately brought down the sitting administration.<sup>19</sup> To offer another example, Soifer (2006) initiated his study of various types of state power in Chile and Peru by skimming a small sample of archival documents. This initial research enabled him to identify patterns worthy of exploration, pin down the time period over which divergences in state power emerged, and determine what questions to ask when considering a larger sample of documents.<sup>20</sup>

### **Case selection**

Information drawn from pre-existing materials – a census, or the full complement of documents surrounding a particular diplomatic crisis, for example – can also help scholars to identify the universe of cases for a particular phenomenon, and can inform case selection at the macro, meso, and micro levels. Skocpol and her coauthors (Skocpol, Marshall, and Munson 2000; Skocpol 2003), in their examination of the nineteenth-century origins of associationalism in the United States, sifted through records from multiple organizations with the goal of building a relatively complete dataset of associations that had enrolled as members 1 percent or more of the American population since the 1830s. In another instance, a scholar studying Supreme Court decisions in India used mentions of particular rulings in newspapers and the secondary literature to create a list of decisions that were “politically important” to the central government. He subsequently vetted that list through a series of interviews with experts, and then returned to print media, reading four newspapers to verify that the cases were indeed politically important in the way he had defined the term in his project.<sup>21</sup> Scholars can also use statistical data they collect in the field to make case-selection decisions. A comparativist studying identity formation in the Balkans sought a sample of schools that varied with regard to the ethnic division of the town in which they were located and by their curriculum. Data on

<sup>19</sup> Example mentioned by a participant in a workshop for this book manuscript held at Indiana University, December 2, 2011.

<sup>20</sup> Personal communication, November, 26, 2010.

<sup>21</sup> Interview, DK-8, August 1, 2012.

population demographics and documents on school organization gathered from various governmental, non-governmental, and international organizations helped her make selection decisions.<sup>22</sup>

### **Conceptualization and measurement**

Pre-existing materials can also help scholars to define their key concepts (or refine their “pre-field” definitions), and develop contextually valid measurement strategies. Cardona’s (2008) analysis of how the institutional design of the Colombian security forces affected the type of armed challenge the government tended to face during the 1946–1954 civil war known as *La Violencia* is a good example. Cardona operationalized and measured institutional design using information drawn from internal documents from the ministries of defense and the interior, press reports from the late nineteenth and early twentieth century, and correspondence from national political figures involved in designing the security forces. Further, he used police personnel records, governors’ correspondence, and documents found in archives in the state of Antioquia to operationalize and measure types of armed challenge to the regime.<sup>23</sup> In order to research the effect of exposure to French repression of a 1947 revolt in Madagascar on current self-reported levels of freedom of expression there, Garcia Ponce and Wantchekon (2011) used recent individual-level survey data, as well as maps and information drawn from archives, to trace insurgency movements and the French reoccupation of the territory.

To offer two additional examples, in her study of opposition to authoritarianism in Syria, Wedeen used such materials as political cartoons and politically oriented television programs – the “forms taken by everyday political contests in Syria” (1999, 87) – to sketch (i.e., to measure) both the contours of resistance and the regime’s attempts to limit and direct the opposition (by selectively tolerating critiques and expressions of resistance in such materials). In another example, Zuern (2011) used cartoons, associational flyers, and social movement posters to examine the evolution of community organizing and the meaning of democracy in South Africa. One cartoon, for instance, vividly depicted the transformation in how local associations conceived of their rights, shifting from portraying smaller figures on the sidelines of a frame submissively asking for a few “breadcrumbs”

<sup>22</sup> Interview, DK-9, August 1, 2012.

<sup>23</sup> Personal correspondence, January 30, 2011.

of local material rights, to a tall figure at the center, triumphantly holding over his head the “whole loaf” of “full political rights” (Zuern 2011, 49).

### Hypothesis testing and generation

Political scientists have also used many types of pre-existing materials to test their hypotheses, refute rival claims, and generate new causal propositions. For instance, in *Useful adversaries* (1996), Christensen advances the argument that the United States and China clashed in the 1950s (in the Korean War and the 1958 Taiwan Straits crisis) because conflict allowed leaders on both sides to mobilize domestic constituencies for larger strategic purposes. He illustrates those mobilization efforts and shows how conflict facilitated them – and refutes competing claims about the roots of the conflict – using information drawn from decades of academic studies of the period in question, opinion polls, memoirs, a range of primary documents from archives and published collections, speeches and telegram communications, and even political cartoons from American and Chinese newspapers. In another example, Jacobs used detailed information concerning decision makers and decision-making processes drawn from documents found in fifteen archives in the United States, Canada, Germany, and Britain to test hypotheses about the considerations that informed policymakers’ and interest groups’ choices during particular historical episodes. Doing so allowed him to show how these individuals’ high-level conceptualization of choices correlated with the kinds of considerations on which they focused (Jacobs 2011).

### Illuminating causal processes and mechanisms

Finally, information drawn from pre-existing materials can help scholars to understand causal mechanisms and processes better – to get at *why* political phenomena are related in a certain way. Archival documents addressing the nitty-gritty of particular political dynamics can be extraordinarily valuable sources for process tracing, for instance. In his book on variation in regime type across five Central American countries in the early twentieth century, Mahoney (2001) showed how critical junctures are choice points, using information drawn from myriad documents (i.e., internal memoranda, laws, decrees, etc.) from multiple presidential administrations to emphasize agency, thus fortifying his path-dependent explanation. More specifically, he used pre-existing materials to identify the dilemmas and strategic choices

that Central American leaders faced during the period of liberal reform in the nineteenth century, and to demonstrate how their choices and strategies led to regime variation the following century.<sup>24</sup> Another scholar sought to demonstrate that state elites adopted new nationalist ideologies in mid-twentieth-century Mexico through the “cooptation of cultural producers” – the integration of intellectuals, artists, and activists formerly associated with the Communist party into the state apparatus and, in particular, onto textbook commissions. He did so by examining meeting documents from those commissions, identifying attendees, and tracing their histories back to the Communist party.<sup>25</sup>

In sum, as these examples demonstrate, pre-existing materials can aid scholars in achieving analytic goals at every stage of research – and thus can contribute to theory building. Of course, as with all data-collection techniques, scholars must reflect critically as they gather pre-existing sources. For instance, as noted previously and discussed further in the next section, information collected from fragmentary pre-existing materials often gains meaning through its combination with information drawn from other sources. Moreover, researchers should consider the implications of allowing pre-existing materials to shape their research questions, the way they conceptualize key phenomena, etc. Choosing questions in this way may make perfect sense when a researcher is focusing on deeply historical topics – particularly if choosing other questions would mean a dearth of data.<sup>26</sup> Yet doing so may lead to the systematic under-study of particular topics, or areas, or groups of people, and thus to skewed and uneven knowledge generation.<sup>27</sup> Scholars can mitigate these concerns by discussing in their published work other questions they *considered* asking and how data-availability put those questions out of bounds. Even with these caveats, we strongly disagree with the contention that collecting pre-existing materials draws scholars into an intellectual vortex that prevents them from seeing the big picture and taking steps toward developing theory to explain crucial political dynamics.

<sup>24</sup> Personal communication, James Mahoney, March 23, 2011.

<sup>25</sup> Interview, DK-2, July 26, 2012.

<sup>26</sup> Indeed, the more accessible data sources are, the easier and more transparent (and thus more replicable) data collection is, and the more easily data sources can be verified by other researchers.

<sup>27</sup> For example, governments in the developing world and elsewhere often lack systematic and reliable statistics on the informal economy, landlessness, illegal money transactions, the role of vulnerable social groups, and illegal migration flows – yet these are topics that are crucial to study (Sadiq 2005, 189).

## Collecting pre-existing materials: challenges and strategies

Field sites hold vast seas of potentially relevant pre-existing materials. This situation presents researchers with a series of challenges: making decisions quickly (and under pressure) about what is important for their project and what is not; finding and collecting the information necessary to answer their questions; and evaluating its evidentiary value. Many of these challenges result from two signature qualities of pre-existing materials: they were generated by someone else for another purpose, and they cannot “talk back.” As one international relations scholar warned, “The archives were not organized to fit your project!”<sup>28</sup> In this section, we offer some strategies to assist scholars with these challenges.

Before doing so, however, we emphasize the importance of beginning to collect pre-existing materials *early* in one’s field stay. There are many reasons to do so. Sometimes these materials are “low-hanging fruit” – publicly available and located in places (e.g., archives and libraries) that are familiar and comfortable for academics. These qualities can make gathering such materials relatively easy, and make doing so a good way for researchers to ease into fieldwork and immediately make progress. Collecting them may be less challenging physically or emotionally than interviewing or beginning to set up an experiment, for instance. Conversely, pre-existing materials that appeared (from one’s home institution) to be relatively safe and easy to collect can turn out to be challenging to gather once one is in the field. Materials that seemed to be publicly available may have an elaborate permission process that can take months to complete. Scholars who postpone collecting supposedly easily accessible information until the end of their research trips may be unpleasantly surprised to discover they have run out of time.

Further, even after scholars verify that data are readily available and easily collectable, the situation can change. Archives can close (or flood, or burn down), and archive staff can go on strike. In what was undoubtedly an unusual run of bad luck, a week after one scholar initiated dissertation research in an archive in Rio de Janeiro, she received a note on her desk indicating that library workers were going on strike; a month later, when they came back to work, the archive closed for inventory; and a few weeks after that, library workers went back on strike.<sup>29</sup> Government ministries, or any

<sup>28</sup> Interview, LM-15, September 10, 2012.

<sup>29</sup> Interview, DK-5, July 31, 2012.

type of organization, can move, entailing a months-long packing up of all documents. And staff can curtail access to certain materials, or remove information they had made available online, without warning. Contacts who have promised access to certain types of information can get fired or become reluctant to provide that access.

Finally, scholars' assumptions about the existence or contents of data sources they postpone collecting can be inaccurate. For instance, Mazzuca and Robinson (2009) planned to develop a new theory of proportional representation in Latin America based on data gathered from historical accounts of electoral reform in the region. When they discovered that no in-depth historical studies existed, however, they had to "become historians" themselves, carrying out archival research and eventually producing two papers: a historical account of power-sharing in late nineteenth- and early twentieth-century Colombia (the "secondary literature" they had originally sought), and an additional piece (drawing on the first) summarizing the cases in light of the theory.<sup>30</sup> The lesson is clear: scholars should identify and determine the accessibility of crucial pre-existing materials early in their field stay, and collect them as soon as they can.

### **Identifying and selecting sources**

While scholars whose data-collection repertoire includes gathering pre-existing materials will want to cast a wide net (Trachtenberg 2006, 162), they will inevitably have to make choices about what to collect.<sup>31</sup> That is, they will need a selection or sampling strategy – a plan for choosing a subset of the available materials.<sup>32</sup> Scholars' choices will generally be guided both by pragmatic considerations based on field conditions, and by the analytic purpose for which they envision using the materials in question. When working in archives, or faced with a library of NGO documents or multiple datasets, scholars should ask themselves: what questions do I want to ask this archive, library, or dataset? To which aspect of my project will the information I gather contribute, and how? How will it help me answer my research question?

<sup>30</sup> Personal communication, November 20, 2010.

<sup>31</sup> While our overall advice is for scholars to focus their searches, series or collections that seem only tangentially related to their topic can contain valuable treasures, and so may be worth a quick skim.

<sup>32</sup> If sources are scattered about various locations, selection choices may need to be made at various levels – location, archive, collection, and materials, for instance.

Beyond these general questions, however, how can scholars decide which pre-existing materials to collect? Trachtenberg describes a theory-guided selection strategy in which scholars start by gathering and analyzing the sources that are the easiest to access, iteratively acquiring a sense of the “architecture” of historical problems, and from there develop an appropriate research strategy and set of questions to ask (2006, 30–50, 140–146, 163–168).<sup>33</sup> One of our interviewees noted that political scientists may have to approach archives in ways that are somewhat different from those of historians:

The historian is often interested in something close to a blow-by-blow account, ordering a series of events or processes chronologically and without gaps, whereas the political scientist is (especially if going in with theoretical hunches) interested in something close to collecting those data relevant to the observable implications of theory. Often the political scientist is having to do the archival work for numerous cases. He or she can't afford to make the same per-case investment. It can be overwhelming to the extent one starts to think too much like a historian about the extensiveness of data you need to collect.<sup>34</sup>

In a similar vein, and particularly in connection with selecting archival materials, Saunders recommends scholars apply what she refers to as a “theoretical filter” in order to identify particular substantive, temporal, or geographic cut-off points.<sup>35</sup> Researchers should be sure not to apply such a filter in a way that would amount to “cherry picking,” however, focusing too much on finding evidence for their reconstruction of an event or their favored hypotheses, and failing to pursue and examine information that would disconfirm their explanation or substantiate rival accounts of the phenomenon of interest.

It bears mentioning that it may sometimes be appropriate, given a scholar’s analytic goals, to assume a position at either extreme of the sampling spectrum. On the one hand, researchers who use pre-existing materials very selectively in their studies – simply to fill in blanks or missing information – may search archives just for particular documents, and thus not need a specific selection strategy at all.<sup>36</sup> On the other hand, even though most scholars will select sources purposively, random sampling from a wider

<sup>33</sup> Trachtenberg’s Appendix II (2006, 217–255) identifies an extensive range of published, semi-published, online, archival, and open sources relevant to international history.

<sup>34</sup> Interview, BR-4, August 9, 2012.

<sup>35</sup> Elizabeth Saunders, 2010, “Archival methods of research,” IQMR, Syracuse University, Syracuse, NY.

<sup>36</sup> Interviews, DK-2, July 26, 2012; DK-7, July 24, 2012.

population of letters, or minutes, or newspaper articles may be possible, and can help scholars resist the “temptation to focus on the strange, the exotic, and the unusual” (see Prior 2003, 150–154, on sampling).

Of course, all sampling decisions are project-specific. Historically oriented social scientists might use various strategies to evaluate secondary sources and select historical accounts as the empirical referents for their studies, for instance.<sup>37</sup> In an archival context, whether to select documents from certain years, related to or written by certain people, or produced in certain contexts, for example, depends on document accessibility, the researcher’s particular intellectual and theoretical goals and how the documents chosen will advance them, and the analytic methods the scholar is employing. The selection choices of a scholar who plans to use information drawn from newspaper articles in her analysis will likewise depend on factors such as the availability of the newspapers (and the form in which they are available, which impacts their searchability), the scholar’s research goals, the techniques she will use to analyze the articles, and the methods she will employ to draw inferences in her overall project.<sup>38</sup>

Two concrete examples are illustrative. When Mahoney was researching Central American political regimes (2001), rather than examining documents from the entire liberal-oligarchic era (approximately forty years from the late nineteenth through the early twentieth century), he focused on the single presidential administration in each country during which the main liberal reforms were passed. And he engaged in “targeted primary source research,” examining the documents on which historians had drawn already, identified by reading the published secondary historical literature on Latin America’s liberal-oligarchic era. This sampling strategy was appropriate because his goal was not to unearth new facts or create new data points, but rather to bring new theoretical principles and a different methodological approach to bear on documents that historians had perused before him, comparing them in a new way to see if he could replicate historians’ inferences while building a broader theory.<sup>39</sup> Another scholar doing archival research on early Latin America who wished to paint a holistic picture of state development chose to examine only government ministries’ annual

<sup>37</sup> For a series of competing strategies, see Goldthorpe (1991, 219–225), Lustick (1996, 615–616), Thies (2002, 362–364), Curthoys (2005, 357–363), and Isacoff (2005); Trachtenberg (2006, 51–78, 199–216) offers a survey of sources of bibliographic information.

<sup>38</sup> See Stockmann (2010, 113–116) on drawing samples from media sources.

<sup>39</sup> Interview, James Mahoney, March 23, 2011.

reports (including statistical data) rather than delving into a plethora of fine-grained documents from multiple ministries.<sup>40</sup>

Scholars are likewise focusing – and placing parameters on – their data collection when they make decisions (or when decisions are made for them) about how to *capture* data.<sup>41</sup> To briefly reiterate our discussion from Chapter 4, scholars might take notes, take digital pictures,<sup>42</sup> photocopy,<sup>43</sup> scan, print from micro-fiche, read (particularly critical) documents into a voice recorder, or some combination of these. Some archives might make their documents available electronically.<sup>44</sup> The important point is that data-capture decisions have analytic implications. Reproducing a source or parts thereof (e.g., scanning or photographing) can often be done comparatively quickly, but will entail less filtering, reduction, and processing. By contrast, when scholars take handwritten notes, they gather information more slowly, but are initiating filtering, reduction, and processing in the act of capture. These tradeoffs are important to consider when scholars face data-capture decisions.

No matter which sampling and data-capturing strategies scholars adopt, they can evaluate the value for their project of any particular source by mapping it back to their data-collection plan. Doing so reveals how pertinent to the project's overall goals the source is, and can help researchers resist what may be a strong temptation to overload. Returning to the principle of flexible discipline, the data-collection plan can help scholars to assess the relevance of information they were searching for and found, *and*

<sup>40</sup> Personal communication, junior comparativist, November 26, 2010.

<sup>41</sup> Not all of the options discussed here will always be available; some archives may not allow cameras, may disallow flash, or may place heavy restrictions on quoting from or copying materials (Hill 1993, 24).

<sup>42</sup> It is not necessary to purchase a fancy digital camera in order to take pictures of documents or newspaper articles; a 10-megapixel camera is sufficient. It can be helpful if the camera has an LCD screen that flips out (so what is being photographed is visible to the seated photographer), and if its battery can be removed. Scholars are advised to buy multiple batteries, as well as a tripod with long extendable legs. They should also determine early on what type of software they will use to organize their pictures. Scholars should practice with all of their equipment and technology at home prior to entering the field to avoid technical difficulties on site.

<sup>43</sup> See Hill (1993, 53) for a compelling set of advantages to photocopies.

<sup>44</sup> Accessing digital editions of newspapers can help scholars to find specific articles, or engage in systematic searches for articles on particular topics using carefully chosen key words. They should consider the reliability of the newspaper's search engine or the external search engine (e.g., Factiva) they are using, and remain mindful of the biases that may be built into search engines developed by for-profit companies. They should also keep in mind that if they search some subset of the sample of newspapers relevant to their study online and another subset manually, their selection technique will differ across the sample, potentially compromising their ability to compare the number or types of articles appearing (which may or may not matter depending upon the contours of their analysis).

of “unanticipated information” that they did not seek or expect to encounter. Moreover, as we have emphasized, scholars should clearly and systematically document and justify all of their decisions concerning data collection. Why were certain documents or articles examined but not others, why were some documents collected while others were not, why were particular documents captured one way and others a different way – and what implications did those choices have for the analysis? Clearly documenting their strategies can help researchers to think clearly about those choices, potentially improving their analysis, and can make it easier for them to solicit advice on the matter. Doing so also makes it easier for scholars working in more than one field site to adapt the sampling, collection, and capture strategies used in one context to another. Keeping track of one’s selection choices and their justifications also facilitates their communication in final written products, thus facilitating transparency.

### **Dealing with missing and elusive information**

Many researchers will face the opposite challenge: not *limiting* their search, but dealing with missing and elusive information. For instance, state records are often incomplete, especially for distant times or marginal populations, and particular archival documents or whole collections may be unavailable.<sup>45</sup> Relevant data may be scattered across many locations. At the document level, sources a scholar is able to access may not yield critical information – a particular person’s name, or the day on which something occurred – making the researcher want to shake the recalcitrant newspaper or document and demand, “Why can’t you tell me *this*!?” Meeting minutes might leave out important comments, and other sources may only be available in “sanitized” or redacted form (Trachtenberg 2006, 157). Indeed, as discussed in the next subsection, the challenge is even greater when one considers that sources may be absent (and existing sources may lack information) that scholars are not even aware they are missing. Such known and unknown silences can be as analytically consequential as they are frustrating.

<sup>45</sup> If the root of the problem seems to be disorganization, one way to “give back” to an archive is to help with organization. One scholar did so at the library at the Museo General de la Policía in Bogotá, an archive managed by one of the institutions he was studying. Parts of the archive were in complete disarray, and one room in particular was full of enormous (approximately 4 feet x 18 inch) ledgers of police personnel records from the 1890s to the 1940s stacked in disorganized piles. This scholar lugged them around, organized them, and catalogued them, leaving the catalog with the archive staff (personal correspondence, January 30, 2011).

Beyond changing their research design, what strategies can scholars use to deal with missing information?<sup>46</sup> First, we suggest that they carefully consider *why* certain sources or bits of information might be missing. The reasons for their absence may imply something of substantive relevance to the project. Moreover, determining why they cannot be located can help in identifying and correcting for any resulting biases, a point to which we return in the next subsection (Trachtenberg 2006, 159). Scholars may also try to determine whether the sources in question are randomly or systematically missing. For example, governments are often legally authorized to impound records for some period of time after they are produced, and place extra controls on or withhold indefinitely particularly sensitive information (Tosh 2000, 43–44). Further, archival collections can be destroyed (unintentionally or intentionally). Under apartheid, the South African state archives methodically destroyed documentary records that could harm the image of the state; today, tattered brown manila envelopes that appear to have once housed large stacks of documents contain nothing more than a slip of paper that says “*Vernietig / Destroyed*” (see Tosh 2000, 42; Pohlandt-McCormick 2005, 299, for more examples). In addition, individuals or organizations may intentionally withhold information from archives so that it is *not* destroyed. Again in the context of apartheid South Africa, activists self-censored, avoiding committing certain types of information to paper or destroying records pre-emptively so they would not fall into the hands of police or security; further, the African National Congress often hid its documents or took them abroad (Pohlandt-McCormick 2005, 300).

Second, it is important to calmly and carefully assess how crucial the missing source or elusive information is for what scholars are trying to measure or evaluate, or the argument they seek to develop. Is it a smoking gun? Can they carry out the analytic task at hand without it? Third, thinking creatively about what substitutions might be valid and whether there is another way to access the information sought is advisable.<sup>47</sup> If the National Archives are not open, or do not seem to contain the relevant documents, might they be found in archives in state or regional capitals? Could important documents concerning ex-colonies be held in archives of former colonial capitals? Maybe other scholars or organizations have reproduced and

<sup>46</sup> For more general discussion of having a “Plan B” for multiple facets of one’s project, see Chapter 3; for guidance on changing analytic strategies or significant aspects of one’s project while in the field, see Chapter 10.

<sup>47</sup> Indeed, where official documents are housed can itself be a revealing piece of data (Polillo 2008, 7).

assembled some or all of the sources for which a researcher is searching. He can also seek to determine whether a related local organization has a clipping service that identified and preserved just the types of newspaper articles he seeks to find. (Scholars who find such a goldmine will need to inquire about the goals that organization was pursuing when selecting articles, and the methodology they employed.) Or perhaps the sources of interest have already been published or made available online.

One scholar of international relations in Europe, upon discovering that a major historical political figure's documents were impounded, turned to the published writings of his press spokesman, who had attended all cabinet meetings and one-on-one sessions in which the figure participated, taking verbatim notes.<sup>48</sup> When another IR scholar could not locate articles on the Indian government's decision-making concerning nuclear weapons, he drew information from diplomatic records found in archives in the United States and the United Kingdom.<sup>49</sup> Of course, even if scholars can obtain *all* the pre-existing materials they seek, they should continually consider how they might triangulate – pairing those materials with information garnered via other data-collection techniques.

### **Identifying bias in and evaluating the evidentiary value of pre-existing materials**

Information drawn from certain types of pre-existing sources may seem more objective than data gathered using more interactive data-collection techniques – perhaps because it is quantitative, for instance, or officially produced. Yet pre-existing sources do not necessarily offer more objective, “factual,” or “uncontaminated” accounts than do interview respondents or focus group participants. The fundamental similarity that characterizes all of the material discussed in this chapter is that it has survived – whether for a few hours, or a few centuries. Surviving sources are almost always a finite subset of the universe of sources relevant to any scholarly research, and the properties of that universe are generally unknown and often unknowable.<sup>50</sup> The available sample (particularly with regard to historical sources) is likely not random, and may well be unrepresentative of the broader whole, and/or

<sup>48</sup> Interview, DK-6, July 24, 2012.

<sup>49</sup> Example given at manuscript workshop, Indiana University, December 2, 2011. See Trachtenberg (2006, 158–159) for an additional example.

<sup>50</sup> Put differently, “only a part of what was observed in the past was remembered by those who observed it; only a part of what was remembered was recorded; and only a part of what was recorded has survived.” The source is Milligan (1979, 185) quoting Gottschalk (1969).

marked by significant bias (Mariampolski and Hughes 1978, 108; Goldthorpe 1991, 213).<sup>51</sup> Many reasons might explain sources' survival, and survival signals neither significance nor objectivity. As political scientist and archival methodologist James Goldgeier has observed, an archival record "is not a truth, it is an artifact."<sup>52</sup>

Different types of pre-existing sources may have different biases.<sup>53</sup> State-generated data, for example, can reflect the agendas and biases of those who created them, and can be oriented and expressed in particular ways for the purposes of communicating with counterparts in other countries, with other state entities, or with society (Chen 2010, 22). Data-gathering agencies may twist or skew statistics – inflate or deflate numbers related to budgets, taxes, personnel, illness, or crime (to give just a few examples), or fail to report certain data – in order to accommodate cultural biases, influence budget allocations, suggest greater state capacity than actually exists (or perhaps lesser capacity, in a bid to attract foreign aid), conceal graft, and for many other reasons. Even if published, statistics must be analyzed and their completeness and reliability evaluated with careful attention to the time periods captured, and consistency in the units of analysis and measurement. "Downstream" sources (containing data that have been processed by the state, e.g., analytical reports or statistical results) may be even more biased than less-processed, raw, "upstream" sources (Chen 2010, 31). Finally, government statistics on phenomena that are more difficult to observe – terrorist networks, the illegal economy, illegal trade, illegal financial transactions, and illegal immigration flows (Sadiq 2005, 182) – are automatically suspect. In short, many factors compromise the availability, validity, and reliability of state-created data.

Other pre-existing sources can have analogous weaknesses. For instance, NGOs' internal documents can be written for funders or other audiences rather than being faithful renditions of organizational priorities or activities. Likewise, even when media offer excellent day-by-day coverage of key events, scholars should be aware of their profiles and ideologies. What implicit or explicit filtering system may be at work sorting stories in and out? Keeping

<sup>51</sup> Record-keeping is often haphazard, and government record-keeping "fickle and political" (Curthoys 2005, 364). Further, particular dynamics or phenomena are sometimes simply much better documented in some periods than in others. For instance, according to one scholar we interviewed, much more was written about nationalism in Mexico after the 1910 revolution (interview, DK-2, July 26, 2012).

<sup>52</sup> IQMR at Syracuse University in June 2010.

<sup>53</sup> The subsequent discussion draws on Sadiq (2005) and personal communication with a senior historian, March 14, 2011; see also Chen (2010).

in mind the goals of the researchers who produced the books or articles a scholar is relying on, and considering what biases those researchers may have brought to their projects and the primary data on which they drew, are also important research tasks.<sup>54</sup>

An additional source of bias in pre-existing materials made available by institutions or organizations (including but not limited to archives) are the multiple routes through which and conditions under which materials arrive there (and fail to do so). Bias can be introduced, intentionally or unwittingly, at every step in the construction and operation of such venues in what Hill, drawing on Schutz (1970–1971, Vol. III) describes as a *sedimentary process*. The destruction and discarding of sources by potential donors, their subsequent choices about what to share from what is left, and the choices made by those who ultimately collect, accept, prepare, and organize the materials and put them on offer, can all introduce bias (Hill 1993, 9–19). Further, the stories of the elite and powerful may dominate such venues, particularly when those who operate them believe materials related to well-known people may lend prestige to their institution. Narratives of traditionally marginalized groups, by contrast, may be underemphasized, stored in a way that makes them more difficult to access, or left out completely.<sup>55</sup> Particularly when such venues function as sites of contention or of “negotiating” the past, or if they have been implicitly or explicitly charged with controlling the memory of particular periods, episodes, or actors, they may reflect and emphasize certain social and political perspectives, thereby “manufacturing pertinence” (Fritzsche 2005, 186), while other aspects of history are excluded (Milligan 2005, 160; Robertson 2005, 71).

The multiple and varied types of bias from which pre-existing materials may suffer inevitably provoke epistemological debates about whether and how to interpret them, draw inferences from them,<sup>56</sup> and evaluate their evidentiary value.<sup>57</sup> Those who take an interpretive approach to analysis

<sup>54</sup> For an account of various types of biases in historical work, see Thies (2002, 359–362); Lustick (1996) recounts how historians’ accounts of the very same events can vary widely due to differences in their personal, methodological, and theoretical commitments, or many other reasons.

<sup>55</sup> The systematic exclusion of certain voices is one rationale for conducting oral history interviews, discussed in Chapter 6.

<sup>56</sup> Goldthorpe (1991) and Dibble (1963) consider the challenges of drawing inferences from pre-existing data, and the former argues that social scientists should not rely exclusively on such data. Shih, Shan, and Liu (2010, 54–57) explore how to create a quantitative dataset from material collected from primary and secondary sources, and Hill (1993, 64–69) also discusses “making sense of” archival data.

<sup>57</sup> For additional useful information on evaluating the evidentiary value of sources and information, see George and Bennett (2004, 97, 99–105, 107–108).

might argue that there is no such thing as objectivity and consequently dismiss efforts to isolate and eliminate bias. More positivist scholars might use various techniques to contextualize their sources and pinpoint the ways in which they are biased, and subsequently seek ways to adjust for any bias identified.

Scholars can begin to assess the objectivity and validity of the information they draw from their sources by carefully scrutinizing them as they gather them, asking what information they might be lacking, and actively seeking to determine what related sources might be missing. They should also investigate the provenance of their sources with an eye to identifying the viewpoints, biases, perspectives, and objectives of their authors or creators. They might ask questions about the founding, history, evolution, reputation, profile, and political orientation of the venue in which they found the sources. To offer a few examples reflecting some of the causes of bias discussed above: was there a particular purpose (commemorative, punitive, esthetic) for constructing the venue? Who gathered or assembled its contents? If the materials were donated, did the donor play a role in the processes or events documented in the archive? What logistical or financial limitations did those who assembled the venue's contents have and how might they have impacted what ended up there? What did the archive compilers expect the uses of the information to be? Is the venue in conflict or cahoots with the government?

An analogous series of questions – generating different types of “external and internal criticism,” in historians’ parlance (Milligan 1979) – may inform scholars’ evaluation of individual sources or groups of sources. When, where, under what conditions, and by whom, was the source produced and/or assembled?<sup>58</sup> What was the producer’s or assembler’s purpose in generating the source and who was the intended audience? What biases might the producer or assembler have had? Who was present when the source was formulated or discussed, and who approved it? As Charles Tilly has suggested, researchers should seek to “interview” the materials they collect in order to develop theories about the generation of evidence and better understand the structure and sources of bias (Chen 2010, 16).

Scholars may also try to assess how representative the “surviving” sources are of any population of which they might have formed part, and how adequately they reflect the broader range of relevant people, experiences,

<sup>58</sup> With regard to a quantitative set of survey data, for instance, scholars might investigate how respondents were sampled and questionnaires administered.

and viewpoints. What other information might be necessary for the researcher to understand the source? How *reliable* is it – how authentic (i.e., not a fake or forgery) and how credible (how likely is it that it accurately transmits the way events occurred, people behaved, etc.)? Was the person who produced the source in a position (geographic, temporal, mental) to know the truth first-hand, and able (socially and physically) and willing to tell it? Is the source consistent or self-contradicting? How well does its meaning accord with other contemporary accounts (Mariampolski and Hughes 1978, 105–110; Goldthorpe 1991, 213; Tosh 2000, 51–62; Thies 2002, 357–359; Trachtenberg 2006, 146–162)?<sup>59</sup> Is the version accessed the only version there is? Was it published?<sup>60</sup>

To be clear, we are not suggesting that materials that do not pass muster on all of these fronts are not useful. On the contrary, their existence, and the particular ways in which they portray, misread, or misrepresent events can be analytically significant. We simply urge scholars to take any biases they encounter into account as they interpret their sources – and to report them in their scholarship so that readers can also take them into consideration.

### Collecting ethically

One might think that ethical concerns, which are often closely related to interaction with human subjects, are less likely to come into play when research entails gathering pre-existing materials. This may in fact be the case. However, no matter in what form they gather information, scholars need to treat that information – particularly if it is sensitive or classified, or brings to light dynamics that have significant emotional weight for those involved – with care, remaining ethically *conscious* through the process of gathering and deploying it in their analysis. Through the book, we emphasize that practicing ethical commitment involves more than following an IRB protocol; rather, it also entails considering the potential effects one's actions, interviews, and observations can have on project participants. Analogously, scholars who collect pre-existing research materials should carefully consider the potential ramifications of collecting (and being known to have received) them even when they have no official designation.

<sup>59</sup> Dibble (1963, 204–210) discusses criteria and rules for evaluating testimony (produced by individuals) and “social bookkeeping” (produced by groups or organizations).

<sup>60</sup> Some historians hold that authors can be more candid when they were not writing for posterity or publication (Tosh 2000, 34); personal communication, senior historian, March 14, 2011; but see Hill (1993, 62–63).

Just because a document has been declassified through some official bureaucratic process does not mean that it is not sensitive, for instance. Particularly (but not only) in less-democratic contexts, there can be a considerable gray area between “explicitly marked secret” and “public,” and scholars should take care to know where the information they are gathering falls on that continuum. Likewise, they should consider what the information’s status implies about how to treat it, about how (and whether) to discuss possessing it, and about what to reveal about how they received it. Read’s study of urban state–society relations, for example, drew in part on internally published neighborhood election records from the Beijing city government (2012, 57–77ff.). Records for the year 2000 were happily provided to the author by city officials, but in later years the same records were deemed more sensitive, even though their official status did not seem to have changed, necessitating acquiring them through an indirect and undisclosed source.

In addition, archives and other venues in which scholars collect information often place limitations on the use or redistribution of their holdings. Constraints are often document- or source-specific – for instance, many documents are under copyright. Accessing and using individual oral history transcripts drawn from a broader collection offer an example: some who provided oral histories might allow all materials to be quoted and used, whereas others may maintain copyright, meaning that “fair use” guidelines apply (i.e., only small portions of the transcript can be quoted or incorporated in another scholar’s work without securing the copyright holder’s permission). Operating ethically means fully adhering to these constraints. At other times, particularly under repressive political regimes, pre-existing materials may have been produced anonymously, with no one claiming ownership. In any of these scenarios, researchers may find themselves navigating political minefields as they seek permission to use, quote, or reprint pre-existing materials in their published work.<sup>61</sup> In short, scholars are just as constrained by ethical standards and considerations when dealing with pre-existing materials as they are when collecting data using other techniques. And even when gathering data involves minimal human *interaction*, their collection can have considerable human *implications* to which scholars should remain attentive.

In sum, understanding as much as possible about their sources and the context in which they were produced, and pairing pre-existing materials with

<sup>61</sup> Interview, LM-25, October 21, 2013.

information gathered through more interactive forms of data collection, can enhance the meaningfulness of all sources collected and improve scholars' ability to deploy them effectively. Collecting data in this way instantiates several of the principles underlying good field research, including engagement with context, critical reflection, triangulation, and ethical commitment.

---

## Preparing for and conducting archival field research

---

While much of the information offered in the previous section is applicable to archival research, a specific set of challenges (and solutions) attend conducting such research in the field. This section considers those issues.<sup>62</sup> Archival research has much to recommend it. The wealth of information archives contain and the level of detail in archival documents can trigger new ways of looking at old issues: archival records can “open a wide window on officialdom,” allowing scholars to “see the state ‘in action,’” glimpsing behind the “velvet curtain” at the “political backstage” (Diamant 2010, 36, 40, 41, 45, 50).<sup>63</sup> Archival work involves “perpetual surprises, intrigues, and apprehensions” (Hill 1993, 6). Archives can be a site for (and can reflect) political, cultural, and socioeconomic struggles and pressures, however, and scholars must unearth them in order to fully understand archives’ contents. Moreover, there are disciplinary biases against archival work (including presentism; see Diamant 2010, 38), and concerns springing from its unpredictable and “necessarily provisional and iterative essence” (Hill 1993, 6).

Of course, archives – and thus archival research – vary significantly. Government archives, university archives, corporate archives, missionary archives, non-governmental or philanthropic organizations’ archives, oral history centers, and collections of papers housed privately<sup>64</sup> differ in terms of rules of access, politics, internal practices, and other parameters. Some archives contain only very old documents, while others (such as newspaper archives) receive new documents each week. Some archives reflect public or private investment in preservation and protection, others do not; some are

<sup>62</sup> Parts of this section draw on modules taught by James Goldgeier and Elizabeth Saunders at IQMR, and presentations at George Washington University’s SICAR.

<sup>63</sup> Diamant (2010) offers an eloquent and compelling discussion of the value of archival work for understanding Chinese politics, and more generally.

<sup>64</sup> In the United States in particular, presidential libraries represent treasure troves of documents and interview transcripts from colleagues, staff, and family of the chief executive. For example, the Kennedy library contains over 1,600 interviews, and the Truman library has more than 500.

well organized while others are in disarray; in some archives, fulfilling requests takes minutes, while in others, it takes months. Generally speaking, in authoritarian regimes (compared with democracies), there may be little presumption that politically relevant information should be available to researchers and more tension surrounding issues of access.

These differences notwithstanding, we believe the strategies discussed here will resonate with and aid most scholars who carry out archival research, no matter what sort of archive they are working in, or how heavily they are relying on the information they find there. We offer two prefatory suggestions. First, as with all data-collection techniques, balance needs to be sought between executing one's pre-determined research strategy on the one hand, and open-mindedness and flexibility – going with what the evidence seems to be indicating, and modifying one's concepts and assumptions as the research unfolds – on the other. Second, scholars should think carefully about what the idiosyncrasies of archival research imply for the rhythm, timing and sequencing of fieldwork. For instance, many scholars who carry out heavily archive-based projects take an initial short “surveying” trip to determine the contents of archives (perhaps developing and deploying a sampling technique) and ultimately revisit archives and reread materials more than once (Hill 1993, 67).

### **Preparation for archival work**

There are several steps researchers can take to prepare for archival work and develop clear goals. Reading the secondary literature on their topic allows scholars to see how debates are structured; scour discussions, footnotes, acknowledgments, and bibliographies for information concerning archives, collections, and documents of potential use; see how archival data can be used to support claims; and begin to envision how they will orient their study vis-à-vis existing work on their topic.<sup>65</sup> Likewise, identifying relevant published collections of interviews and primary documents and examining their introductions and editorial notes to see what archives and sources are mentioned can be very useful. Further, scholars can begin to build their “archive network,” contacting the authors of secondary sources, scholars cited therein, and other researchers who have carried out similar studies to ascertain what archives and collections they researched, how they did so,

<sup>65</sup> Trachtenberg (2006, 51–58, 199–216) provides useful discussions of how to systematically identify relevant work; Hill (1993, 33–36) offers thoughts on identifying promising archives.

and whether they might be open to communicating about or sharing any of their documents. Once in the field, scholars can expand their networks to include in-country experts (Polillo 2008, 7). To give one example of the value of contacting local academics, the scholar of Mexican nationalism mentioned earlier noted that the idea of looking at notes from textbook commission meetings (to trace the evolution of official nationalist ideas) came to him through talking with historians with years of experience in the education ministry's archive.<sup>66</sup> The same can be done with journalists. Taking these steps can do a great deal to orient scholars and their work. For instance, a researcher working in Europe explained that footnotes from relevant secondary sources, published compilations of archive documents, and conversations with experts at local institutions played an essential role in helping focus his search in actual archives: "Without that I would have been completely overwhelmed."<sup>67</sup>

Researchers should also learn all they can about archival methodology – about systematically searching for relevant sources, and about drawing inferences from documents. They can do so by examining political science scholarship based on archival research, as well as the social science literature focused on archival methods.<sup>68</sup> Engaging with the (admittedly sparse) literature and debates from history and historiography can also help social scientists to develop the tools that they will need to read and interpret different types of sources: a different lens is required to analyze policy memos, advertisements, posters, and lists of donors, to name just a few types of sources.<sup>69</sup> Researchers can also attend institutes and workshops,<sup>70</sup> and possibly audit courses on archival research in the history department at their home university. And they might scan web sites that can introduce them

<sup>66</sup> Interview, DK-2, July 26, 2012.      <sup>67</sup> Interview, BR-4, August 9, 2012.

<sup>68</sup> With regard to books, Hill (1993) is an excellent, practical introductory primer; Prior (2003) discusses how to use documents in social science research; Trachtenberg (2006) provides a practical guide to studying international history; and Frisch *et al.* (2012) is likely the most extensive "how-to" source in political science (despite focusing on American politics), helpfully integrating the insights and perspectives of archivists. Mahoney and Villegas (2007) discuss how primary sources (and secondary sources) can be used, and how inferences can be drawn from them.

<sup>69</sup> See, e.g., Tosh (2000). Burton (2005) has assembled a fascinating collection of essays on archival research; in particular, Curthoys (2005) provides a helpful discussion of interpreting historical documents.

<sup>70</sup> SICAR and IQMR were both mentioned previously; other opportunities include the Methods Cafe sponsored by APSA's Interpretive Methodologies and Methods conference group and the short course on archival methodology, both offered during the annual APSA meeting. More specialized workshops also occur; for instance, in November 2010, Stuart Shulman ran a workshop on the US Freedom of Information Act (FOIA), examining electronic content management, e-discovery, redaction, and other topics at the US Department of Agriculture (also available via Webinar).

to archival research: the London School of Economics' "Archives made easy,"<sup>71</sup> H-Net Discussion Networks in general (and H-Diplo Resources in particular, especially "Guides to doing research" under "Organizations"),<sup>72</sup> and the American Historical Association's web site.<sup>73</sup> As they learn more about interpreting sources, scholars may wish to put what they are learning into practice by, for instance, developing draft coding sheets (on which they can get local input once in the field).<sup>74</sup>

As their target archives begin to come into focus, scholars should do what they can to increase their familiarity with them. Beyond learning about their history and evolution, scholars can tap the "archive network" they have begun to create to learn about their target archives' culture, pace, and working environment (e.g., accommodations, clientele, noise, foot-traffic, copier use);<sup>75</sup> their social practices and informal rules (for instance, dress code);<sup>76</sup> and what can be brought into the archive (technology in particular).<sup>77</sup> They can also inquire about how helpful the archivists are and what local customs there might be for working with them (should they bring the archivist a gift of some sort?) and seek to identify other tips for operating in the archives. Scholars might cast a broad net – contacting faculty or graduate students from various disciplines, in the United States or elsewhere – and of course should continue to develop their archive network once in the field. Where better to find people who know an archive inside and out than in and around the archive itself?<sup>78</sup>

Preparation also entails investigating target archives' holdings (determining what exclusions there may be) and internal logic (i.e., the way materials are classified, organized, and ordered), determining the number of items in potentially interesting collections, and identifying what related collections might be useful. Ascertaining what finding aids (brief statements about the scope, structure, and/or contents of a particular collection) are available is also important. To do so, scholars may be able to consult archival inventories

<sup>71</sup> [www.archivesmadeeasy.org](http://www.archivesmadeeasy.org). <sup>72</sup> [www.h-net.org/~diplo/resources](http://www.h-net.org/~diplo/resources).

<sup>73</sup> [www.historians.org/info/research.cfm](http://www.historians.org/info/research.cfm). <sup>74</sup> Interview, DK-5, July 31, 2012.

<sup>75</sup> Personal communication, January 30, 2011.

<sup>76</sup> In particular, they might check whether the use of ultra-thin gloves to protect the documents is required.

<sup>77</sup> One scholar working in Brazil, for example, was allowed to bring nothing more than paper and pencil into one archive in which she (consequently) spent months in Rio (telephone communication, DK-5, July 31, 2012).

<sup>78</sup> It can even be useful to ask others who work on one's topic or a closely related one – but who do not seem to have used the archive or accessed the documents in which one is interested – why they did not use the archive or collections in question, or whether they know anything about the archive or know anyone who has used it.

and catalogues: some archives publish these on the web or in electronic or print book form. For archives with paper catalogues that can only be accessed from the archive (or with no catalogues at all), researchers might contact scholars or others familiar with the archive, or archive personnel, to seek information about archival holdings. If targeted archives' holdings are in a foreign language, scholars should evaluate their language skills and assess the advisability of using an interpreter.

Another step involved in preparing for archival research (one with implications for fieldwork timing and sequencing) is determining in what *format* (e.g., paper, microfilm, fiche, etc.) the materials of interest are available. Paper can take more time for scholars, archivists, or intermediaries to retrieve; it is also more delicate and may thus take longer to read and process in the archive. As a rule of thumb, the older one's sources, the more fragile (and, likely, less organized) they will be. Microfilm/fiche may be easier to retrieve and handle, but have unique capture challenges: printing from reading machines, if it is possible at all, often requires plunking in coins or using a payment card for each image, which is time consuming. The more familiar scholars become with target archives in advance of beginning work there – perhaps initially during an exploratory trip – the better they can assess how relevant the archives' resources are to their project, how much time research there should take, and which documents to review in what order.<sup>79</sup>

A final aspect of preparation is emotional. While collecting data in the manner under discussion here can be extremely rewarding (and less nerve-wracking than certain types of interviewing or ethnographic work), it can be challenging emotionally – isolating, lonely, and cerebral. Such emotional challenges can, in turn, make it more difficult to deal with the ambiguity fieldwork entails and to summon the patience and persistence required. Both the up-sides and down-sides of archival research need to be planned for in advance.<sup>80</sup> Scholars are encouraged to check log-books to see who else is working in the archive, what topic they are working on, and what documents they are using, and to talk to and befriend others researching there. Beyond the potential for companionship and rich intellectual exchange, such people may have learned some of the archive's tricks and secrets, may have other

<sup>79</sup> Hill (1993, 27–39) discusses name-oriented and other archival search strategies.

<sup>80</sup> Concerning archival research's (admittedly minor) physical health risks: scholars may be lifting heavy boxes and sitting for extended periods of time. Those with allergies should consider their exposure to dust in the archives.

relevant information to share, and may be interested in coordinating on requesting and using documents.

### **Navigating archival access and intra-archival relations**

Gaining access to archives and their collections often has both practical and interpersonal aspects. To begin with the practical, scholars should check what restrictions there may be on accessing the archives they wish to visit; verify whether they will be open on the days they plan to visit (keeping national holidays in mind); and confirm their working hours. In addition, they should arrive at their field site with their credentials fully in order. They should determine what paperwork, identification, documentation, permissions (from the country and/or the archive), and/or letters from their home institution are required to gain access to the archive.<sup>81</sup> Investigating whether they need to enter the country on a particular type of visa or register with the archive as a reader is also recommended.

Learning about the terms and process for gaining access to the specific documents they wish to consult is also important; special permission from depositors or a government official, for instance, may be required to view certain collections. Moreover, scholars might be able to arrange in advance for some documents to be ready for their perusal upon their arrival. Indeed, they might need to request them ahead of time; in this case, they will need to know how to place their request (by mail, phone, or internet) and how far in advance to do so. They might also ascertain whether there are any limits on how many items can be requested and examined per day, and whether any access fees are charged.<sup>82</sup>

It can also be advisable to develop good relationships with archivists (and any other intermediaries who retrieve the sources scholars wish to use) for several reasons. Archivists may have the ultimate say on which materials researchers can access. Accordingly, whether archivists look favorably upon a researcher, and whether they are sympathetic to and interested in her project, can make a big difference to her archival experience. For example,

<sup>81</sup> Scholars ought to bring several copies of any such letter, with requisite official marks, with them to the field, and perhaps versions translated into the local language.

<sup>82</sup> A further note on requesting materials: when intermediaries are retrieving documents, it can be beneficial to give them broad enough requests such that the scholar is doing the sorting and actual choosing of sources rather than the intermediary. Researchers should also be sure to keep track of what they have requested, rather than relying on archive staff to consistently retrieve (all of) the correct materials – and should not necessarily shy away from requesting closed files or sensitive documents (at least once or twice) in case an archivist or intermediary might be willing to share them.

one interview respondent recounted making a point of learning enough of a local language in Senegal to greet the archivist every day – a strategy that paid off when this scholar was the only visitor allowed to work through lunch and to exceed limits on the distribution of materials.<sup>83</sup> Another suggested that the relationships and trust she built with the clerks at the institution where she was doing research – which she suggested she did by relying heavily on them rather than immediately “going over their heads” to speak with senior staff – resulted in her gaining access to materials other scholars had been unsuccessful in finding.<sup>84</sup> Moreover, archivists who view themselves as the protector of their collections may be reluctant to offer access to scholars they suspect will write something negative based on the materials they wish to access.<sup>85</sup> One of our interview respondents described having just this experience in a German archive: staff members assumed, from a connection the researcher mentioned having with another scholar, that the researcher was not sympathetic to their point of view, and were consequently less forthcoming with materials.<sup>86</sup> Finally, archivists often have intimate knowledge of the collections under their care (including collections whose existence may not be public) and specialized tools to search them (perhaps using finding aids not on the shelves). They may also be willing to introduce researchers who are working on similar topics to each other.

Of course, scholars should not take what archivists say as gospel, nor expect them to know everything about archival collections. Archivists can be powerful political figures in their own right, and may have incentives to portray their institution or its holdings in a certain way, or may have been trained to approach the collections under their care from a particular perspective (e.g., relatively conventional or politically safe). When making recommendations, they may (intentionally or unintentionally) overlook documents in which researchers (with their own perspectives and goals) might be interested, or portray them in a way different from that in which researchers would see them.

These caveats notwithstanding, given the influence archivists can have over one’s research experience, identifying connections one might have with archivists (perhaps through one’s host institution or an interview respondent), and building relationships and trust with them (perhaps inquiring with

<sup>83</sup> Interview, LM-7, September 20, 2012.

<sup>84</sup> Interview, DK-17, August 24, 2012.

<sup>85</sup> Additionally, one never knows when one may need to contact archivists or intermediaries after leaving the field to request missing documents, ask that a particular document be re-photocopied, etc. In this vein, see Hill (1993, 56–57) on “archiving at a distance.”

<sup>86</sup> Interview LM-4, August 27, 2012.

others who have worked in the archive about how to do so), can be vital for facilitating entrée to archives and to specific collections. We offer three concrete suggestions for gaining archivists' favor and trust.

First, early on in their research, scholars should read archivists' introductory essays in collection inventories. These can reveal a great deal about archivists, the contents of collections, and what related documents may be available elsewhere. Second, particularly if their study is politically sensitive, scholars should seek to present themselves, their topic, and their project in a neutral or even positive light (for instance, in any "orientation interview" they have with the archivist: Hill 1993, 41–44). Using uncontroversial language when describing the research, and keeping in mind how archive staff unfamiliar with the ways of foreign academics might interpret their goals and behavior, are both good ideas. For example, a scholar studying welfare largesse with regard to the elderly (and at the expense of other population segments) in Italy might present the project as one focusing on Italy's generous support to the aging. Finally, scholars can reassure archivists that they will use information from the archive responsibly, and perhaps even agree to keep certain pieces of information confidential (as appropriate for their study).

### **Organization and analysis**

Given the sea of information in which archival researchers swim, remaining organized is crucial.<sup>87</sup> Of particular importance, as we discuss elsewhere (see Chapters 3, 4, and 10), is devising a personally appealing organization system that is readily searchable, and from which the scholar can easily retrieve items. Particularly if he anticipates sharing his data or releasing information for replication, the system should be easily understood by others as well. A critical facet of a scholar's organizing system should be a log including every folder extracted from an archive box (whether or not the document in the folder was collected), perhaps using ATLAS TI (or Excel or Word). Such logs can be annotated with descriptive information about the documents and ideas for how they relate to the broader analysis.<sup>88</sup> Further, a system for highlighting items that look potentially useful (or less useful) and justifying

<sup>87</sup> We are grateful to Elizabeth Saunders for the vast majority of the suggestions offered in this subsection, which she discusses in the presentation on "Archival methods of research" that she gives annually at IQMR.

<sup>88</sup> Scholars might consider the following naming protocol for documents: Place/Date/ArchivistName/Series/Subseries/Box/Folder/Document Title.

those judgments could be created. All of these annotations are preliminary steps toward analysis. For those items scholars capture (in hard-copy form or digitally), it is also crucial to keep bibliographic information with the data themselves (see Hill 1993, 69). When photocopying or scanning documents, for example, scholars can copy or scan the title page and/or copyright page and keep it with the other individual pages reproduced; an alternative is to put a slip of paper with the bibliographic information on the copier glass when copying. Analogous processes can be used when taking photos of documents. Alternatively or in addition, scholars can take a picture of the label of the box from which the files that housed the photographed documents were drawn, making sure to capture the folder tab in the picture.

In terms of organizing data (scans of documents, digital pictures, coding sheets, etc.) on their computer, scholars might impose a file/folder structure that mimics the archive's organization, and then organize all of the images they are capturing and other materials they are gathering by folder. Scholars might consider using an image management system such as Picasa to store and display pictures; Picasa adopts the file structure on one's computer and allows the cropping, straightening, rotating, brightening, and whitening of images. Further, scholars can create captions in Picasa (great quotes from particular documents can be added right into the caption field) and tag documents with key words or ideas – and then search document photos by keyword or quote. Ideally, data will also be organized and stored in a way that will be recognizable by computers in the future. However, to reiterate, the most important criterion is that the organization system be intuitive and easy for the researcher to use so his documents remain quickly accessible to him. And as we have also repeatedly emphasized, scholars should back everything up – their framework documents, their logs, their notes concerning sampling, and all their data – by burning them onto a CD, copying them onto a secondary hard drive, or uploading them to a network or cloud storage site. When those options are unavailable, they can email their materials to themselves or a trusted colleague.

Most importantly, scholars should strive to maximize the intellectual value of the organization process by beginning to process and analyze data as they organize them. In doing so, they should consider their broader analytic goals. For instance, if they ultimately hope to conduct discourse or content analysis on the materials they are collecting, that may dictate a particular way of organizing them. They may wish to get all data into one format (perhaps by creating digital pictures of articles for which a hard copy was created in the archive). Fully transcribing documents that contain large amounts of highly

salient information for the project,<sup>89</sup> or making a “table of contents” of some documents, noting the page and line number where the most crucial bits of information about particular topics can be found, can be helpful. Scholars should also continue taking notes on, describing, and summarizing their documents and further annotating their log. More broadly, seeking to connect the items they have collected to their data-collection plan and their timeline (if they have created one; see Chapter 3) can be very useful. In sum, developing and deploying a strategy for collecting and analyzing data *in tandem* represents the most efficient route to the research goal. We offer further thoughts on how to do so in Chapter 10.

---

## Conclusion

---

Gathering pre-existing materials in the field is an important mode of data collection in which practically every scholar who conducts fieldwork engages, even if sporadically. While collecting pre-existing sources is often equated with archival research, this chapter sought to demonstrate that materials that can prove critical to one’s research can be of infinite types, and can be found in locations far beyond archives. We strongly encourage scholars not to overlook potential sources of useful knowledge, and to think openly and creatively about the data-collection enterprise. We also compared and contrasted pre-existing materials and the process of collecting them with more interactive forms of gathering data. When working with extant sources, scholars have less influence over what information is potentially available, and thus over what data they can ultimately generate. They are essentially collecting bread crumbs left by earlier actors and scholars who had their own interests and agenda in mind when they produced or collected those sources, and by the archivists who made them available. Despite these differences, a steady commitment to ethical practice is no less important when collecting and analyzing pre-existing sources than it is when engaging in more interactive forms of data collection. Moreover, because of their differences and the distinct types of data they produce, *combining* less-interactive and more-interactive modes of data collection is an excellent strategy for enhancing inference. For certain kinds of questions and projects, opportunities for triangulation among sources of multiple types abound.

<sup>89</sup> Another option, if photos of documents are very clear, is to run them through an Optical Character Recognition (OCR) program to convert them to text.

This chapter also discussed strategies for making theoretically informed and methodologically sound decisions about which pre-existing materials to consult and what information to collect, and for appraising sources' evidentiary value. At the root of many of our suggestions is the importance of engaging with the immediate context in which pre-existing materials were found, learning as much as possible about the more distant context in which they were created, and critically reflecting on how those contexts likely affected the contours, content, tone, and, indeed, resilience of the pre-existing materials a scholar has collected to use in her analysis. We also examined how deploying such materials can help scholars accomplish multiple analytic goals and piece empirical puzzles together.

Throughout, we emphasized the benefits, for the researcher and for readers of her work, of practicing transparency when working with pre-existing materials, another research principle we emphasize. Scholars make myriad choices when collecting pre-existing sources, particularly in an archival context. Documenting those choices as they are made – clearly explaining how and why a certain sampling technique was adopted, or particular archive eschewed – will help scholars remain clear on how the data underlying their conclusions and inferences were compiled. Conveying the details of those choices in their published work will allow readers to develop the same confidence in scholars' conclusions as the scholars themselves have.

Collecting pre-existing materials in the field can be exciting and rewarding – particularly when one happens upon a dataset that had proven elusive, a “smoking gun” document, a bumper sticker with a priceless quote, or a political cartoon that perfectly captures the gist or irony of the political dynamic under study. When considering this form of inquiry, we urge scholars *not* to allow their minds to conjure up stereotypical images of being locked away in the damp, drab depths of a musty archive slowly slogging through heaps of paper by candlelight. As one interviewee put it:

It's possible to trace policy without field research. But connecting the dots in official records reflects countless assumptions about what happened and why ... Digging through archives can test those assumptions, allows you to learn what was in people's minds, what pressures they were under. For me, it's completely exciting.<sup>90</sup>

<sup>90</sup> Interview, BR-9, August 16, 2012.

In sum, playing detective – filling out the complex empirical story of a cataclysmic historical event, or a quickly developing international showdown, step-by-step and piece-by-piece from documents written by those who actually sat around the bargaining table, or from daily newspaper accounts – can be exhilarating and profoundly gratifying.

## Interviews, oral histories, and focus groups

While not all political scientists consider gathering pre-existing materials to be a core mode of data collection, interviewing has traditionally been an essential technique for gathering data in the field.<sup>1</sup> With the exception of the site-intensive methods discussed in the next chapter, employing other data-collection techniques rarely allows scholars to achieve a level of dynamism or fluidity of interaction with those whose behavior they hope to understand equal to that which can be realized in an interview. Interviews allow us to gather information to generate detailed, holistic descriptions, capture varying perspectives, discuss processes, unearth competing interpretations of events, identify the micro-foundations of macro-patterns, and frame hypotheses (Weiss 1994, 9–11). Moreover, while political actors' formal decision-making processes are more visible and available than ever before in today's information-rich environment, interviews allow us to explore the informal interactions and behaviors that can be equally important to political outcomes (Beckmann and Hall 2013, 297).

Indeed, according to the results of our survey of US political science faculty, interviewing is the most common form of data collection for political scientists who engage in fieldwork. Fully 81 percent of field research projects reported in the survey made significant use of interviews – the highest percentage associated with a data-collection technique about which we asked. Specialists in all subfields use interviews liberally: 92 percent of field research projects undertaken by comparativists involved them, and 84 percent and 71 percent of IR specialists' and Americanists' projects did so, respectively. Even 50 percent of projects reported by political theorists involved interviews. Oral history figured in 31 percent of all projects (and was thus

<sup>1</sup> See Dexter (1970, 13), Kvale and Brinkmann (2009, 7–14), and Platt (2012) on the history of interviewing in the social sciences and its emergence as a more important data-collection technique in political science as part of the behavioral revolution and with the greater emphasis on empirical methods of investigation in the mid twentieth century.

employed in a higher percentage of projects than was survey research), and focus groups figured in 13 percent. The frequency with which scholars combine interviews with other data-collection techniques demonstrates that political scientists across the methodological spectrum conduct interviews – even scholars whose main focus is experiments, or who spend most of their days buried deep in a library poring over archival documents.

Despite the significant advantages and prevalence of interviewing, few faculty or graduate students have any formal training in or preparation for conducting interviews in advance of their field research. In fact, graduate students and faculty of all ranks told us, sometimes sheepishly, that they conducted their very first interview only after they had arrived in their field site. As we noted earlier, political science graduate programs tend to focus methodological training on how to *analyze* data rather than on how to *collect* them. Our interview data also suggest that even the qualitative methods or field research methods courses offered in some graduate programs generally dedicate only a couple of hours to the discussion (and sometimes the practice) of interviewing. The assumption seems to be that interviewing requires a skill set that everyone has: if you can talk, you can do an interview.<sup>2</sup> Relatedly, there is relatively little *written* on interviewing in political science; a few recent article symposia (Leech 2002b; Ortals and Rincker 2009), an intriguing manual penned by political scientist Lewis Anthony Dexter (1970) more than forty years ago, and a recent edited volume (Mosley 2013a) are important exceptions.<sup>3</sup>

Yet while interviewing may look natural or easy to an outside observer, it involves much more than sparking a conversation between a researcher and an interlocutor. Further, scholars differ on what that “much more” entails. More positivist researchers may attempt to remain as neutral as possible while interviewing, hoping that by keeping themselves “out of the data” they will elicit relatively objective approximations of empirical truth that they can use as evidence in their analysis. In contrast, a more interpretive researcher might insist on the subjective nature of the interview itself, understanding it as a vehicle for revealing constructed representations of various potentially competing truth claims. Many field researchers seem to adopt elements from both traditions, as others have also discussed (Soss 2006). No matter how

<sup>2</sup> Mosley, in the Introduction to her edited volume on interview research, recalls her advisor’s suggestion about interviewing: “Just talk to people” (2013b, 1).

<sup>3</sup> Other political scientists who have written about interviewing include Weiner (1964), Aberbach, Chesney, and Rockman (1975), and Rathbun (2008).

scholars lean epistemologically and think about the production of knowledge, however, conducting interviews effectively – and feeling comfortable doing so – involves extensive preparation and practice.<sup>4</sup> There is a great deal scholars can and should learn about gaining access to potential respondents, structuring interview protocols, formulating and asking questions, and gaining respondents' trust, to name just a few topics. Lack of preparation in these and other areas can have negative consequences for the quality of the data scholars collect as well as for continued access to respondents.<sup>5</sup>

As with other data-collection techniques discussed in this book, the models and practices of interviewing have been examined in a voluminous literature in anthropology, sociology, and psychology in particular (Adler and Adler 2001; Gubrium and Holstein 2002; Leech 2002b; Kvale 2008; Rubin and Rubin 2004, 104–5). Our aim is not to replace the existing literature – indeed, no book, let alone a single chapter, could. Instead, we hope to build on it by highlighting how the unexpected contingencies of working in the field may influence the conduct of interviews and the data gained from them. Hence, rather than discussing interviewing in its neat, abstract form, we try to reveal the messy reality of interview research in the field, illustrating how field conditions, and the limits and tradeoffs (and opportunities) they create, affect the conduct of interview research. While our discussion concerns face-to-face in-person interviews, some of the advice and lessons we offer are also relevant for telephone, voice-over-internet, or email interviews.

In the chapter's next section, we introduce briefly three types of interviewing: (1) in-depth interviews with individuals; (2) oral histories exploring a person's memories of the past; and (3) focus groups aimed at uncovering different viewpoints within a group on a set of topics. We highlight their differences in terms of the number of participants, the role of the researcher, and their degree of structure, and consider their strengths and weaknesses. We also discuss combining interviews with other data-collection techniques, showing how triangulating adds analytic value. In the chapter's third section, we draw on scholarship from all subfields of political science to demonstrate

<sup>4</sup> Kvale and Brinkmann (2009, 47–60) offer a clear and concise discussion of different epistemological takes on interviewing.

<sup>5</sup> In what he describes as a "small editorial aside" in an article footnote, Goldstein (2002, 672) makes a rather sweeping claim about the deleterious consequences of inadequate preparation for long-term access, writing that: "Our discipline's access to elites in Washington, especially members of Congress, has been hurt by massive amounts of poorly trained students and scholars being unprepared for interviews."

how interviewing provides compelling data that can contribute to every stage of theory building – helping scholars to identify a research question, select cases, refine conceptualization and measurement, generate new hypotheses, and understand causal mechanisms and processes.<sup>6</sup> In the fourth section, we discuss some of the issues that researchers face when interviewing in the field, and share strategies for addressing common challenges. While no single solution will address the array of challenges that scholars interviewing in the field may confront, preparing in advance but remaining flexible in practice, engaging with the cultural and political context of the field site, and critically reflecting on what is being learned – being what Giovanni Sartori referred to long ago as “conscious thinkers” (1970, 1033) – will help guide scholars to effective responses to each unique dilemma. Our suggestions in this final substantive section, then, underline several of the principles that we contend underlie good political science field research.

## Types of interviews

We distinguish among and compare three types of interviewing: in-depth interviews with individuals, oral histories, and focus groups.<sup>7</sup> We organize our discussion according to the interview process rather than the status or expertise of the interviewee (elite vs. non-elite). Of course, researchers face divergent challenges and opportunities when they interview respondents from different social classes and who wield different levels of political power, and they may use information gathered through those interviews for different analytic purposes.<sup>8</sup> Nonetheless, with regard to operating in the field, the most important differences among interviews concern the number of individuals involved, the role the researcher plays, and the structure of the interaction (see Table 6.1). Although the boundaries between interview types are fuzzy (one technique can shade into another as researchers employ them in the field), each has characteristic advantages and drawbacks. As we also discuss, they can often be profitably combined – and mixed with other forms of data collection.

<sup>6</sup> For a similar discussion focused specifically on in-depth interviewing, see Lynch (2013).

<sup>7</sup> See Morse (2012, 194–197) and Kvale and Brinkmann (2009, 143–160) for finer-grained distinctions and characterizations of interview types.

<sup>8</sup> Scholars have attempted to clarify that in-depth interviews are not exclusively with wealthy or powerful actors, but confusion persists. See Leech (2002b) who draws on Dexter (1970) to discuss what is meant by elite interviewing; see also Dexter’s (1970, 3, 18) distinction between “standardized” and “nonstandardized” interviewing.

**Table 6.1** Three types of interviewing

|                                |                                   | In-depth interviews  | Oral history interviews   | Focus group interviews   |
|--------------------------------|-----------------------------------|--|---|--|
| <b>Number of participants</b>  |                                   | 1–2  | 1   | 6–10   |
| <b>Role of researchers</b>     |                                   | Provocateur  | Active listener   | Facilitator  |
| <b>Interview constitution</b>  | <b>Degree of structure</b>        | Varies from loose to highly structured; less structured than a survey    | Low; usually researcher identifies starting point and participant guides narrative flow | Moderate; researcher/facilitator poses questions but order and emphasis shaped by participants |
|                                | <b>Question type and ordering</b> | Generally open-ended; questions asked in variable order                  | Open-ended; narrative proceeds chronologically  | Prepared list of open-ended questions; order and emphasis shaped by group                      |
| <b>Potential advantages</b>    |                                   | Can reveal unique perspectives and sensitive, highly charged information | Can reveal unique historical perspective and causal processes                           | Can reveal tensions and disagreements  |
| <b>Potential disadvantages</b> |                                   | Specificity of viewpoints; inefficiency                                  | Particularity; representation as unscientific   | Potential for group think  |

### In-depth interviews

When researchers say they “do interviews,” they are often referring to in-depth interviews. These can of course vary in formality – our understanding includes a formal appointment with a diplomat arranged months in advance, an informal or even spontaneous meeting with an academic colleague in the field, and everything in between. Such interviews are usually conducted with one individual. On occasion, there will be two interview subjects, often with associated or conjoined sets of expertise, for example, a husband and wife, or two business partners. At other times, the respondent may have a trusted aide he prefers to have present, or an official spokesperson, such as the *okyeame* for Akan chiefs in Ghana (Yankah 1995).<sup>9</sup> Likewise, researchers generally carry out in-depth interviews by themselves, although they may be accompanied by a research assistant or a translator (Fujii 2013). While the researcher may be disinclined to “count” aides and translators as interview participants, every

<sup>9</sup> If the number of respondents increases beyond two, the dynamics may shift toward a group interaction more akin to a focus group, described below.

person present, even someone who quietly observes just for a brief moment, can fundamentally shape an interview and the data produced.

The role of the researcher in an in-depth interview may vary depending on how structured an exchange is planned, the nature of the research question, the stage of the project at which he finds himself, his interview objectives, his epistemological perspective, the characteristics of the interviewing context, and the respondent. Nonetheless, in general, the researcher acts as provocateur. He poses stimulating questions, listening actively as he probes and moves the interview forward to cover an array of topics or questions. In-depth interviews are often characterized as conversational, but the interviewee usually spends more time talking than the researcher (a balance referred to as "proportionality").

Greater variation exists with regard to how actively scholars direct the exchange – or redirect it when the respondent goes off in a new direction (an attribute referred to as "directionality"). Some in-depth interviews are highly structured: scholars enter the interview with a precise list of questions prepared in advance to be asked in a particular order.<sup>10</sup> Scholars with these sorts of protocols intervene frequently to guide the conversation, motivate the respondent, and keep the interview focused on their questions. Other interviews are extremely loose, with the order and substance of questions almost entirely guided by the respondent. Interpretive researchers who emphasize the subjectivity of the interview encounter tend to conduct less structured interviews in which topics and themes of importance are identified and constructed through the interaction, and unanticipated detours are the main focus of the analysis. Perhaps most often, in-depth interviews will be somewhere in between. In one form of semi-structured interviewing, for instance, the investigator arrives with a list of "information needed" or "questions to get to," but allows the conversation to unfold in a relatively flexible fashion. Indeed, most researchers seek to keep their minds open to what they can learn when a respondent departs from their agenda – for instance, to illuminate dynamics or questions she thinks are important, or to indicate how she frames debates. Indeed, the tendency toward a lower degree of structure is one of the critical ways in which interviewing differs from survey research. Where the guiding principle of survey interviewing is to

<sup>10</sup> Such an approach, which results in interviews being more similar in question content and order, can produce data that are more comparable and allow for hypothesis testing and the quantitative analysis of interview responses (Leech 2002a, 665).

precisely replicate question wording and order across respondents, in-depth interviewing allows for customization to the individual respondent.

With regard to the questions asked in in-depth interviews, no matter how structured the overall exchange, questions are usually open-ended.<sup>11</sup> However, interview questions can vary greatly with regard to how they are formulated and what information they seek. Sometimes scholars ask vague questions, relying heavily upon the interviewee to choose and construct the information conveyed. At other times, researchers formulate queries with the goal of eliciting concrete information about specific events, dynamics, or phenomena, or responses that speak to particular conceptual categories or will help them to measure a particular phenomenon. Of course, given the vagaries of the field context, an interview that was envisioned to be structured may, in the end, become more semi-structured or free flowing.

In-depth interviews can help scholars to gain a sense of social context, to probe how citizens understand the political world, and to derive relevant categories of analysis. They can be particularly well suited for identifying and detailing elite political actors' unique experiences, perspectives, or viewpoints. No other individual will have quite the same perspective as the Minister of Defense for Israel during the conflict with Lebanon in 2006, or the leader of a political party when it made a fundamental change in its platform. In-depth interviews can also be particularly valuable when the topic under study is highly controversial, sensitive, or politically charged, meaning that potential respondents may be reluctant to discuss the matter publicly or in a larger group. Moreover, scholars who carry out interviews can potentially gain more in-depth "meta-data" about their respondents as compared with those who conduct surveys: rather than simply noting that a particular picture or a book was present in the respondent's office or that staff were gathered just before the interview, they can inquire about these things as part of the encounter. What they learn can help scholars to develop additional insight into interviewees' behavior and to draw better inferences and conclusions from their data. Other writings further

<sup>11</sup> Closed questions provide pre-determined answers from which respondents choose, limiting their latitude. For excellent discussions of the relative merits of open-ended and closed interview questions, and criteria scholars can use to decide which approach to take, see Aberbach *et al.* (1975, 3–8), Schuman and Presser (1981, 79–112), Aberbach and Rockman (2002, 674), Rivera *et al.* (2002, 686), Beckmann and Hall (2013), and Leech *et al.* (2013). Consider also the use of open-ended questions *together with* a brief questionnaire with closed questions in Rivera *et al.* (2002, 686) and Beckmann and Hall (2013), or in tandem with a lengthy questionnaire as in Stokes *et al.* (2013).

discuss benefits of interviewing compared with survey research (Soss 2006; Leech *et al.* 2013, 209–211).

Interviewing has some identifiable disadvantages, however. Respondents sometimes offer self-serving and distorted accounts. An interview subject might convincingly portray her view to be authoritative and objective, or might suggest that it is uncontested, even when it is biased or controversial. Also, like oral histories, any one interview provides a narrow viewpoint on the phenomenon of interest.<sup>12</sup> And, in the words of one of our own interview respondents – one who advocates conducting interviews in tandem with other data-collection techniques – “done well and reliably,” interviews are “expensive in time, money, and thought [and] . . . inefficient.”<sup>13</sup>

### Oral histories

Oral histories are more often associated with the disciplines of history, and, to a lesser extent, anthropology and sociology (White, Miescher, and Cohen 2001; Charlton, Myers, and Sharpless 2006; Giles-Vernick 2006; Atkinson 2012) than with political science. Nonetheless, even when directed by historians, oral history projects often focus on collecting histories from important political leaders or on important political events.<sup>14</sup> As their name suggests, oral histories are more exclusively focused on the past than are in-depth interviews: they are an oral record of an individual’s life experiences (also known as “life histories”), or her memories of an event, or a community’s, organization’s, or family’s history (to give just a few examples). They might be targeted or quite sweeping in scope. The respondent may have been a direct eyewitness to the historical occurrence or time period of interest, or may be conveying knowledge of a past beyond her own lifetime or generation, perhaps recounting memories and stories passed down from older generations.<sup>15</sup> In essence, the personal accounts obtained through oral history research are generally used to crosscheck or supplement incomplete or biased information available in existing written records – that is, to triangulate.

<sup>12</sup> Rubin and Rubin (2004, 47–48) and Dexter (1970) offer some criteria to help scholars evaluate the appropriateness of interviewing as a data-collection technique for their project.

<sup>13</sup> Interview, DK-6, July 31, 2012; see also interview, LM-18, September 14, 2012.

<sup>14</sup> For instance, the University of Virginia’s Presidential Oral History Program, which includes both political scientists and historians, is working to debrief central political figures in several late-twentieth-century presidential administrations. Dexter (1970, 93–99) offers some additional examples.

<sup>15</sup> We do not make a distinction here, as some historians do, between eyewitness accounts as “oral history” (Perks and Thomson 1998; Thompson 2000) and stories passed down between generations as “oral tradition” (Vansina 1965, 1985).

Oral history, as a research technique, has an important political history itself.<sup>16</sup> In the years following World War II, oral histories frequently focused on the private narratives of political elites, particularly high-level government leaders. Beginning in the 1960s, historians began to popularize oral history as a way to obtain evidence about the past in places where historical events were not recorded in writing, or where the written record systematically excluded the voices of less powerful groups. This movement to study “history from below” paralleled temporally the civil rights and women’s rights movements in the United States and the decolonization of much of Africa and Asia.<sup>17</sup> In more recent decades, political scientists and others have begun to use oral histories from above and below. For instance, they have been used to reveal American policymakers’ behind-the-scenes perspectives,<sup>18</sup> and to uncover village-level politics and gender dynamics in the process of revolutionary change in China (Seybolt 1996; Hershatter 2011). Unfortunately, political scientists’ use of oral history interviewing techniques is seldom acknowledged explicitly in their published work, raising the question of whether scholars could and should engage more directly with the rich literature from the discipline of history on how to conduct this type of interviewing effectively.

Like in-depth interviews, oral histories are usually conducted with one individual at a time. Even when the objective is to learn the history of a particular community or culture, generally scholars conduct a series of individual interviews. When people are speaking in a group, tensions can arise between different narratives and representations of history that can unproductively squelch a particular rendition. Yet the principal idea motivating this technique is that individuals have unique memories and historical understandings of the same dynamics, which merit independent investigation. In view of the strong focus on the individual, it is especially critical for

<sup>16</sup> The American historian Allan Nevins founded the first oral history project at Columbia University in the late 1940s. The Columbia Center for Oral History has one of the largest archives of oral history, with over 8,000 audio and video transcripts of interviews. This institution continues to play a leading role in the field, particularly in manuscript management and oral history training. See <http://library.columbia.edu/locations/ccoh.html>.

<sup>17</sup> The expansion of oral history projects may have been particularly central to the development of African-American studies, gender studies (Bornat and Diamond 2007), and African history (White *et al.* 2001) during this period. Prominent examples of oral history are the work of Theodore Rosengarten (1975), a historian who drew on more than 60 hours of oral history interviews to write about the life of an African-American sharecropper who joined a union and resisted the racial oppression of the South in the 1930s; and Jan Vansina (1965, 1985), who used oral tradition to reveal and legitimize the precolonial history of communities in Central Africa.

<sup>18</sup> For example, on how to respond to Russia after the Cold War (Goldgeier and McFaul 2003).

the researcher to learn in advance all he can about the interviewee and how he or she fits into the historical and political milieu. The researcher can then ask searching questions that will elicit details and nuances of the respondent's perspective and how it contrasts with others'.

Oral histories tend to be less structured, more conversational, and more respondent-driven than in-depth or focus group interviews. The researcher frequently identifies the starting point of an oral history interview by using a known event as an anchor to stimulate the recall of the respondent; he may even present the respondent with relevant documents or photographs for her consideration. As the interview progresses, the researcher may gently probe in an effort to shape the interview or keep the respondent focused on the topics of interest, ask additional open-ended questions, or encourage the interviewee to double back and offer more details on particular subjects. But overall, compared with in-depth and focus group interviews, the interviewer often takes more of a back seat: he provides the time and space for the respondent to tell her story, allows her to drive the chronological narrative, and plays the role of active listener while the respondent guides the flow of the discussion. A researcher who has used oral history interviews to gather information about traumatic events put it this way: "I issue an invitation to hear a narrative, and I speak very little" – while remaining curious, engaged, and respectful.<sup>19</sup> Given the heavily respondent-centric nature of oral histories, the researcher often ends up spending considerable time with each interviewee. Due to the effort and emotional energy respondents often must invest in order to recount a personal narrative from the past, oral histories are frequently conducted during multiple encounters.

Oral history interviews can provide unique historical perspectives on political events for which other sources of information are unavailable – perhaps because the society is closed, due to national security strictures, because written accounts or records were not kept during the time period of interest or have been lost or destroyed, because certain groups' stories and perspectives are excluded from written sources due to discrimination, and so on. For example, MacLean was only able to capture the local Ghanaian experience of the Great Depression using oral histories with elder village residents; the written sources in the British archives were penned exclusively from the viewpoint of British colonial officials in the Gold Coast (2010, 107–108).

<sup>19</sup> Interview, DK-19, August 27, 2012.

A potential disadvantage for political scientists of employing this valuable data-collection technique is its alleged “low scientific value” due to the extreme particularity of the data it produces, and its representation by some as more “folklore” than truth.<sup>20</sup> Some scholars may find it hard to imagine that theoretically interesting or generalizable claims could be constructed on the basis of the fine-grained, personalistic information collected through oral history interviewing.<sup>21</sup> Political scientists might meet such objections by shaping their oral history interviews to address targeted theoretical concerns, interviewing multiple subjects, interviewing the same subject multiple times,<sup>22</sup> and perhaps combining oral history with other forms of evidence in a process of triangulation. Carefully describing the conduct of oral histories in one’s scholarship and thereby contextualizing the information gained through them – that is, being transparent, a principle of good field research – could also help to enhance the credibility of evidence gathered through this powerful data-collection technique.

### **Focus groups**

In a focus group interview, a set of people are invited to participate in a group discussion on a pre-determined topic or set of issues, and are queried concerning their ideas, views, perceptions, or attitudes. Unlike oral history interviewing in which scholars seek to isolate particular individuals’ narratives and representations of history, the goal with a focus group is precisely to generate interaction among participants. While this interviewing technique was initially developed by sociologists in the late 1940s, it is most frequently associated with marketing research and advertising campaigns, which began to use focus groups in the 1950s. Focus groups regained some popularity with social scientists in the late 1970s and into the 1980s, although political

<sup>20</sup> One indicator of this viewpoint comes in the form of IRB practices: in 2003, the Oral History Association and the American History Association obtained agreement from the US government to exempt most oral history interviewing from IRB review. Nevertheless, egregious violations of human subjects in other disciplines have provoked increasing vigilance by many campus IRBs, meaning that many projects using oral history are still required to pass some level of local IRB review (Townsend and Belli 2004; Shopes 2007; Schrag 2009). See also [www.oralhistory.org/about/do-oral-history/oral-history-and-irb-review](http://www.oralhistory.org/about/do-oral-history/oral-history-and-irb-review).

<sup>21</sup> In a classic essay reflecting on the early maturation of oral history, Grele (1978) responded to critiques that oral histories were at times only collecting unimportant trivia (Tuchman 1996), noting the importance of their being theoretically grounded and methodologically rigorous.

<sup>22</sup> Grele (1998) argues that oral history deepens our understanding of history more than reading the written record does precisely because those who carry out oral histories are interacting with their sources and can return repeatedly to ask them to recount particular events, or explain in more detail.

scientists only began to employ the technique in greater numbers during the 1990s (Copsey 2008). Just as political scientists used focus groups more and more, so too political parties, political campaigns, and incumbent political administrations increasingly employed the technique to uncover attitudes toward particular initiatives, public policies, and campaign strategies (Savigny 2007).

The most obvious distinction between focus groups on the one hand, and in-depth and oral history interviews on the other, is the greater number of participants simultaneously involved in the former. As such, focus groups may seem to augment the efficiency of interviewing, allowing researchers to speak with more individuals. Indeed, this impression may have contributed to the increased use of focus groups over time,<sup>23</sup> given evolving constraints on funding for field research. Yet, the purpose of focus groups is to obtain *group* data, not data associated with particular individuals. Researchers closely monitor how the substance of individuals' comments evolves through the course of intra-group interaction and discussion. Scholars debate the ideal size for a focus group. If the group is too large, individual participants may not have ample opportunity to share their opinions, and may be inclined to withdraw or hold side conversations. Conversely, if a group is too small, it can be difficult for group dynamics and interactions to emerge. While field conditions may recommend a larger or smaller group, typically the most effective focus groups range from six to twelve members (Krueger 1994, 27–28) and last between 1 and 2 hours.

The optimal degree of heterogeneity and familiarity among focus group members is also a source of scholarly disagreement. Ultimately, a researcher's question and hypotheses will drive focus group recruitment strategies (Short 2006, 104). For example, if the researcher wishes to know how political beliefs and political participation vary among generations, she might organize several focus groups of different age cohorts. The field-site context may also affect the extent and type of heterogeneity that is productive for group interaction (see Hennick 2007). In some field sites, women and men would not be combined in one group. In other contexts, mixing individuals of different citizenship statuses or ethnicities within a focus group would hinder the rapport necessary for candid conversation. Researchers also need to consider whether familiarity among focus group participants would increase or decrease their willingness to communicate and divulge information. While

<sup>23</sup> Our survey results bolster this point: while only 6 percent of projects reported on that were carried out in the 1960s and 1970s employed focus groups, 15 percent of projects carried out from 2000 on did so.

in some contexts it would be impossible or undesirable to assemble a group of total strangers, in other contexts a group of participants without pre-determined conceptions of each other and with little probability of encountering each other after the focus group would generate better data and enhance the prospects of maintaining collective confidentiality.

Focus groups vary with regard to their degree of structure. Some resemble highly structured group interviews with the researcher's standard set of questions providing the framework for eliciting information (Hertel, Singer, and Van Cott 2009, 307). In others, the researcher poses relatively general or abstract questions about the topic, but the order and time spent on each question are heavily shaped by the participants. In either scenario, the researcher's role is primarily to facilitate the group's interactions (Greenbaum 2000), creating a safe and open environment, without judgment, in which participants feel comfortable sharing their individual opinions and personal experiences. Leading a focus group is as much an art as a science. The researcher must have extensive knowledge of the political and social context in order to anticipate and respond to the verbal and physical cues of trust, obfuscation, withdrawal, reticence, and anger (Wellings, Branigan, and Mitchell 2000) – and to manage conflict, domination,<sup>24</sup> and, importantly, silence. The researcher's goals and contextual conditions also influence whether and how she guides the discussion in and out of controversial terrain. Because of the challenges of facilitating, observing, and recording the conversation, scholars usually conduct focus groups with the assistance of at least one other person. As always, the presence of assistants will affect the data gathered, and researchers should consider the implications of their choice of assistant.

As noted previously, focus group interviewing is uniquely positioned to reveal how individuals interact in a public group setting, exchanging views and influencing each other's perspectives; researchers gain insights both by listening to the conversation and by watching the group encounter. Focus groups can expose the range of ideas, attitudes, norms, or opinions concerning an event, topic, or issue; elucidate the tensions and points of

<sup>24</sup> One researcher who frequently uses focus groups prior to developing a survey explained how his study team typically dealt with a domineering focus group participant. One member of the team would approach the person, explain that they noticed that they had "really important things to say," and draw the person away from the group to participate instead in an individual interview (interview, LM-10, September 18, 2012). The literature on focus groups cited earlier provides many additional techniques for dealing with a range of challenging focus group participants. See also [www.psfieldresearch.org](http://www.psfieldresearch.org) for a generic handbook that can be adapted for facilitator training.

disagreement among them; and demonstrate how group dynamics shape them. They are particularly useful when an event or issue is contested – when a productive tension can emerge as the focus group proceeds – yet not so charged that individuals are unwilling to speak freely in a larger group. Indeed, while an interviewee's viewpoint may go unquestioned in an in-depth interview or oral history, in a focus group one participant can challenge another or point out inconsistencies. Thus under the right conditions, conducting a focus group allows a researcher to triangulate in real time, gathering multiple perspectives. Focus groups can also be a relatively efficient strategy for collecting information about a community or an organization (e.g., when it was founded, how many members it has, and so on).<sup>25</sup> In all these ways, focus group interviews differ from more particularistic in-depth interviews and oral histories, in which respondents necessarily frame responses in terms of their own characteristics and perspectives (Short 2006, 104–105).

Yet focus groups have some downsides as well. It can take an enormous amount of time to schedule and organize a focus group, as doing so multiplies the access and logistical challenges posed by individual interviewing (discussed later in the chapter). More substantively, participants' opinions can change in the presence of others (Krueger 1994),<sup>26</sup> raising the specter that focus groups may produce "group think," in which dissident opinions are squelched by the tendency to give way to the majority view (Michell 1999).<sup>27</sup> Focus groups may also become dominated by a few strongly opinionated participants (Mosley 2013b, 7). Indeed, social hierarchies may influence respondents' interactions, proliferating positionality issues (Hertel *et al.* 2009, 307). The focus group setting also increases confidentiality concerns: participants might be unwilling to provide information in the presence of a "group of witnesses" if they believe doing so could come back to haunt them "professionally, politically, or personally" (Dexter 1970, 5). All of these challenges can affect the data that are gathered via focus groups and their evidentiary value. By assembling groups carefully, fostering a welcoming atmosphere, and outlining ground rules, researchers can minimize the possibility of tyrannical consensus, or, at the very least, carefully

<sup>25</sup> Interview, LM-10, September 18, 2012.

<sup>26</sup> Of course, these shifts themselves, and the moment at which, way in which, and direction in which they occur, can be important data points.

<sup>27</sup> Copsey (2008) argues that "group think" is less likely for political issues that are salient, and where participants have more firmly held beliefs and opinions, than for preferences about non-political consumer products, for instance.

observe it (Short 2006). And, of course, researchers who notice any of these dynamics should indicate in their notes – and ultimately in their write-up – what occurred, why, and what effects it might have had on the data they collected.

By discussing these different types of interview techniques in turn, we in no way mean to suggest that they are mutually exclusive – quite the opposite. Understanding the differences and complementarities among them, and their advantages and disadvantages, helps researchers choose which one (or more) to use in their projects. For example, focus group interviews can reveal disagreements that might be examined through in-depth interviews, or deep-rooted foundations for a contemporary conflict that could be explored through oral histories. Just as productive can be using interviews in tandem with other types of data collection, to which we turn next.

### **Combining interviews with other data-collection techniques**

Interviewing can be very profitably combined with other kinds of data-collection techniques, and often is. Interviewing may complement site-intensive methods (see Chapter 7), for instance, if ethnographic observations raise critical questions the researcher can then address individually with those whom he is observing. Alternatively, scholars can use written sources and journalistic accounts to contextualize oral history narratives (White 1995; Giles-Vernick 2006, 92–93),<sup>28</sup> to fill in gaps when respondents' memories fail, or to cross-check information conveyed orally. Likewise, information gained through interviewing can help to corroborate or fill gaps in information from written sources or formal accounts, identify different points of view, garner first-hand and behind-the-scenes knowledge, and learn about underlying dynamics. Lynch (2006), for example, combined interviews with archival research in order to plug holes in the archival record and inquire about particular policy actors' motivations. Learning from interviews with Chinese citizens about ongoing changes in state policy helped one of our interview respondents to identify trends and patterns in government statistics on welfare recipients.<sup>29</sup>

Scholars can also conduct interviews and focus groups while designing a survey (Knodel 1997) in order to see whether concepts resonate with respondents, test-drive language, and “get a sense of the opinions, outlooks, or cognitive maps of people who are similar to the research subjects” to

<sup>28</sup> Interview LM-7, September 20, 2012.

<sup>29</sup> Interview, DK-18, August 24, 2012.

whom the survey will ultimately be administered. Doing so can inform the themes, question order, language, and pre-specified answers to closed-ended questions included in survey questionnaires (Lynch 2013, 37–38). Rogers (2013) conducted follow-up interviews with survey respondents in an effort to evaluate whether they were understanding questions evoking concepts used in previous survey research on racial group identity and political behavior in the way he intended; when he determined they were not, he conducted more interviews to pinpoint whether the problem related to question clarity or broader conceptual issues. Interviewing can also help scholars establish the universe of cases for a project and develop sampling frames for survey research (Lynch 2013, 34–35). Finally, interviewing can be combined with field experiments. In-depth interviews or focus groups can suggest causal explanations to be tested in experiments; scholars may also interview experiment participants in order to debrief them and to gain information to help them interpret their experimental data.<sup>30</sup>

## How interviewing builds theory

Data gained through interviewing can contribute to achieving many analytic goals including (1) formulating research questions, (2) selecting cases, (3) conceptualizing and measuring key variables, (4) generating potential explanations, and (5) illuminating causal processes and mechanisms, and thus help build theory in multiple ways. Of course, researchers operating from different epistemes think about the process and products of interviewing differently, and will thus vary in exactly *how* they use interview data to build theory. We hope our discussion demonstrates that interview data can be analytically useful for scholars regardless of their epistemological commitments.

### Formulating and refining research questions

Interviewing can play a critical role in one of the most fundamental stages of any project: the formulation and refinement of the research question. Interviews can be invaluable for double checking that the assumptions that undergird the framing of a research question are correct, and for evaluating whether one's project is original, compelling, and possible to execute.

<sup>30</sup> See Chapter 9 for examples.

Accordingly, many researchers carry out informational individual or focus group interviews in which they explore these issues during a preliminary visit, or soon after beginning to conduct fieldwork. Open-ended interviewing can reveal political dynamics or perspectives or illuminate new problems and avenues of inquiry that are not readily apparent in the published literature (and thus, potentially undetectable prior to arriving in the field). For instance, based on interviews she conducted with guerrillas, one of our interviewees expanded her project to examine the role of the media with respect to a broader range of political activism than she had originally anticipated studying.<sup>31</sup>

### **Case selection**

Data sources accessible from a scholar's home institution may not allow her to identify the appropriate universe or relevant sample of cases. Interviews can help scholars to select cases for study that will allow them to test the most promising hypotheses and answer their research question. For example, Baumgartner and his collaborators on the Advocacy and Public Policymaking project, which examines the process of lobbying and policy-making in Washington, DC, sought to identify for analysis the policy issues on which lobbyists were active. They used a database of lobbying reports filed with the US Senate to construct a sampling frame of organizations involved in lobbying, and then drew a weighted random sample of 100 of those organizations. They interviewed a staff member at each organization sampled, asking him to identify the issue he had worked on most recently. The issues those staffers identified became the sample of 98 policy issues (Goldstein 2002, 670; Leech *et al.* 2013).<sup>32</sup>

An interview respondent who was studying how policymakers and members of the scientific community understand and discuss health inequalities sought to perform content analysis on proposed legislation that could affect health inequality. She thus faced the challenge of identifying legislation that had been proposed but never passed. She addressed this challenge through interviews – speaking with epidemiologists and policymakers to identify all legislation proposed (including which ministries might have created draft proposals), in what committees it was developed, where and

<sup>31</sup> Interview, DK-5, July 31, 2012.

<sup>32</sup> See also the Advocacy and Public Policymaking project's web site at <http://lobby.la.psu.edu>.

when it came up, etc. These interviews helped her to create the sampling frame from which she ultimately selected legislation to study.<sup>33</sup>

### Conceptualization and measurement

Interviews can also contribute to theory by helping scholars conceptualize and develop strategies to measure key phenomena, and evaluate how cases score on key variables. First, interviewing can aid researchers in exploring emergent concepts, further developing concepts that have not been fully theorized in the literature, or understanding how several concepts are related and linked. It can also help them to identify a concept's contested nature – for example, if interview respondents implicitly or explicitly suggest a range of understandings connected with a particular term (Gallie 1956). Interviews can also reveal the existence of a relatively consensual “on-the-ground” meaning of a particular term that challenges a well-established scholarly conceptualization. For instance, Schaffer (1998) explored everyday understandings of “democracy” in Senegal through “ordinary language” interviews.<sup>34</sup> For his respondents, democracy had little to do with civil liberties and the exercise of political rights through competitive elections; instead, Senegalese spoke of democracy in terms of solidaristic voting, social cohesion, and collective economic security.

Interviews may also help scholars to refine conventional conceptualizations (for instance, by identifying a concept's key subtypes or dimensions), or to challenge standard typologies. For example, through an extensive series of in-depth interviews with current and former leaders of squatter settlements in Lima, Peru, Collier identified several types of settlement formation that had not been examined in the literature on the topic (Collier 1976, 138–139). For another researcher, the information gathered through interviews with former leaders of Brazilian parties that collapsed between 1985 and 1995 and with politicians who had changed parties in the 1980s and 1990s, considered in the context of his knowledge of Chile, Argentina, and Uruguay, led him to realize that “level of institutionalization” was a critical element missing from Sartori's typology of party systems based on the number of parties and polarization of the system.<sup>35</sup>

<sup>33</sup> Interview, DK-15, August 21, 2012.

<sup>34</sup> Schaffer's book presents a discussion of the theoretical underpinning and method of “ordinary language interviewing.”

<sup>35</sup> Interview, DK-7, August 1, 2012.

Information gained in interviews can also help scholars refine their approach to measurement, no matter whether they are engaging in qualitative or quantitative analysis (Lynch 2013, 34–38). For instance, one of our interview respondents created a new measure of “wheelchair access” using focus groups. He and a co-author asked focus group participants (wheelchair users) to highlight the items that were most important to them on a Department of Justice checklist containing hundreds of requirements for making facilities wheelchair-accessible. Using respondents’ input, they devised a measure of “wheelchair accessible” rooted in the wheelchair community, and thus, they believed, more valid.<sup>36</sup> Interviewing can also help scholars to evaluate how well measures of a concept developed in one context work in another.

Relatedly, interviews can be used to acquire data to measure key concepts in a study – to score cases on important variables. Most straightforwardly, interview respondents can identify sources for data scholars can use to measure their central concepts. Yet interviewees can also offer information that directly aids in measurement. Brinks (2008), for instance, aimed to explain variation in the conviction rate for police homicides across Argentina, Brazil, and Uruguay. One aspect of his argument focused on victims’ socioeconomic status. However, he met challenges locating the data to score victims on that variable since not all governments published complete police homicide statistics, and because such incidents were not always reported in local newspapers. Hence, Brinks relied heavily on interviewing to determine victims’ social class: he asked respondents who had known a victim what kind of job he had held, and how much education he had received, and requested that they place him on a class scale.

### **Generating hypotheses**

Interview respondents’ answers to well-crafted questions – and information they provide completely unprompted – can illuminate important causal factors, aiding scholars in generating or developing their arguments. Most basically, researchers may simply ask their interlocutors why a phenomenon occurred as it did, in the hope that their accounts suggest plausible hypotheses.<sup>37</sup> Indeed, sometimes interviewees’ responses can be directly converted

<sup>36</sup> Interview, DK-14, August 10, 2012.

<sup>37</sup> Of course, respondents’ causal interpretations need to be considered with the same rigor and skepticism as any hypotheses the researcher may have developed. See Beckmann and Hall (2013,

into an argument. One of our respondents recounted how an explanation he developed – that open-list proportional representation generates personalistic campaigning – was “direct fed” to him in interviews.<sup>38</sup> Yet interviews can also be used to generate or refine hypotheses in more subtle ways.

For example, interviews can highlight informal exchanges and interactions that have significant implications for important political phenomena. Beckmann (2010) draws on interviews with White House and Senate staffers to uncover informal networks and contacting, and to test the operational tenets of presidential coalition building on Capitol Hill. Lessing (2012) generated hypotheses to explain the interaction between drug trafficking organizations and the state in Mexico, Brazil, and Colombia through interviews with police, army, and government officials, among other strategies. Leebaw’s (2011) interviews with various actors involved in the transition from apartheid in South Africa helped her generate and develop her argument that the creation and evolution of the Truth and Reconciliation Commission (TRC) were heavily shaped by a critique of legalism – negative perceptions of the uses and capacity of law rooted in the legalization of violence and human rights abuses under apartheid. Soss explains how long interviews with US welfare recipients helped him to understand how their political action or quiescence resulted, in part, from their understandings of themselves and of their status in relationship to state agencies (2000, 2006).

### **Illuminating causal processes and mechanisms**

---

Interviewing can also be a crucial strategy for uncovering causal processes and identifying causal mechanisms. For instance, in her study of the high levels of civic engagement by American veterans of World War II, Mettler (2005) used data gained through interviews (as well as a survey) to argue persuasively that both the higher levels of education received by those who took advantage of the GI Bill, *and* their personal experience as participants in the program, led to more active civic participation by the “Greatest Generation.” While the nuance might have been missed had the researcher relied

197–198) and Leech *et al.* (2013, 219) for cautions about adopting this line of inquiry in interviews; a scholar interviewed in connection with this project also expressed deep reservations about this type of question (interview, DK-6, July 31, 2012).

<sup>38</sup> Interview, DK-7, August 1, 2012.

solely on the survey data, in interviews veterans frequently described the GI Bill as an inclusive and fair program that treated them like “first-class citizens.” In another example, in her study of Chinese workers’ propensity to use the legal system to resolve workplace disputes, Gallagher found significant differences between younger and older workers. However, her interviews revealed that the divergence was tightly connected with the major political and economic changes experienced by the older workers through their lifetimes: the causal mechanism at work was not simply demographic, but deeply political (2013).

In sum, interviews can help researchers to accomplish a variety of analytic goals at every stage of inquiry, and ultimately can help them to build theory. Yet interviewing can be challenging in countless ways. The next section explores some obstacles researchers might encounter when interviewing in the field, and offers strategies to overcome them.

---

## Addressing the challenges of interviewing in the field

---

Throughout this book we strongly advocate, and demonstrate the importance of, carefully preparing for fieldwork. The sociology, anthropology, psychology, and emerging political science literatures on interviewing offer a wealth of practical advice on conducting interviews and analyzing the information they produce, and we strongly urge scholars planning to use this data-collection technique to consult those literatures. In this section we offer strategies to help researchers decide when and whom to interview, construct effective protocols or interview guides, gain access to respondents, carry out and capture interviews, and evaluate the evidentiary value of interview data. We highlight how field conditions can inform research design, interview preparation, and interview conduct – thus shaping what scholars will be able to claim on the basis of their data. Much of the advice we offer is equally applicable to interviews with individuals, oral histories, and focus groups, and throughout the section we seek to highlight issues related to, and strategies that might be particularly useful for, a specific type of interviewing.

At the outset, we want to re-emphasize a point made in Chapter 4: interviewing is not for everyone. It requires a special type of patience, polite persistence, and perseverance, as well as a relatively outgoing personality. One senior scholar suggested that when he was a doctoral student, his advisor had “sent him” to the field on an exploratory trip, an experience

that taught him he would *not* be an effective interviewer.<sup>39</sup> Scholars should consciously consider their own needs, as well as those of their project, as they make decisions about which techniques will allow them to collect the data needed to answer their questions.

### When and whom to interview

Two linked questions concern how to sequence interviews, and how to choose respondents. Many researchers do some purely informational or exploratory interviews when they arrive at a field site to familiarize themselves with the new context, inquire about available resources, etc. Depending on the amount of time a scholar has available, his question, and the particularities of the field site,<sup>40</sup> it can be advisable to postpone conducting formal interviews with high-profile political actors until he has firmed up his baseline knowledge of the topic through interviewing academics, journalists, and others who *study* the phenomenon, and scouring available written sources.<sup>41</sup> These “softer,” often semi-structured, interviews can help researchers to identify which actors to interview, assess data availability, refine interview questions and construct better protocols and interview guides, and acquire basic information that will help them maximize the payoff from later interviews with actors who may only be available once.

Concerning specific respondents,<sup>42</sup> there is no clear metric to identify the “right” interviewees – to ensure that the researcher has not interviewed an “anomalous slice of a larger population, which has distinct central tendencies in its attitudes and behaviors” (Cammett 2013, 142).<sup>43</sup> Whom to interview – and what sampling or selection strategy to employ – will depend heavily on scholars’ interviewing objectives, practical considerations, and the field

<sup>39</sup> Manuscript review session, Syracuse University, June 2011.

<sup>40</sup> For instance, Boas’s (2009) study of electoral campaigning required that he interview and videotape politicians prior to national elections, leaving him less latitude to delay beginning to interview.

<sup>41</sup> Interview, LM-6, August 30, 2012. We emphasize again the need for flexible discipline, however: important political actors can get busy or travel, for instance. Moreover, interviews sometimes materialize out of the blue: a researcher might be interviewing a (very connected) academic, and on a basis of a phone call made by the latter, presto, the former has an interview with the Minister of the Interior in 30 minutes. Taking advantage of such serendipity is often advisable. Balancing comfort, knowledge, and pragmatism is an ongoing challenge when deciding when to conduct interviews (interview, LM-12, September 6, 2012).

<sup>42</sup> We offer additional general thoughts on case selection (at multiple levels) in Chapter 3.

<sup>43</sup> There is likewise no clear metric to help scholars to determine the optimal number of interviews to conduct as the issue is project-dependent. Bleich and Pekkanen (2013, 90–91) suggest a “saturation” model; see also Chapter 10.

context itself. There is a great deal written on sampling and selecting interview respondents (in both the general methodological literature and the literature on interviewing and focus groups), and we strongly urge scholars to consult these resources.<sup>44</sup> Here we offer just a few rules of thumb.

Researchers carry out interviews for different purposes. Among other possibilities, they might be hoping to discover particular pieces of information, obtain certain insights, or understand particular dynamics; they might be using interviews to inform or shape work mainly based on other data; or they might wish to make generalizations from a sample of people interviewed to a larger population of which they form part (Goldstein 2002, 669). In the first and second scenarios, in which scholars are likely using interview data to *develop* causal explanations, they might seek to interview outliers, exemplars, individuals with specialized expertise on a particular topic, actors who can inform them about particular events or outcomes, or those whose voices are missing from accounts gained through other sources (Martin 2013). They will often wish to speak with people connected in multiple ways to, and with varied perspectives on, the issues and questions that are central to their study – in order to avoid becoming convinced by unchallenged narratives or getting labeled as partial to one side of a controversial issue (Dexter 1970, 43).<sup>45</sup> In these cases, some form of non-probability sampling will be optimal to identify interview respondents.

For instance, scholars might engage in quota sampling (in which they set proportions so a sample of interview subjects includes certain segments of the population), purposive sampling (basing selection on characteristics of respondents that are relevant to the analysis), chain-referral or snowball sampling (allowing an initial set of respondents to suggest further respondents, who suggest further contacts, and so on),<sup>46</sup> or even convenience sampling (interviewing those who are available and agree to be interviewed).<sup>47</sup> One less formal way to sample that somewhat combines these techniques is to identify and list, as exhaustively as possible, the scholars

<sup>44</sup> For excellent discussions of choosing respondents and of sampling in connection with interviewing, see Weiss (1994, 15–33); Rubin and Rubin (2004, 64–70); Tansey (2007); Beitin (2012, 243–247); Beckmann and Hall (2013, 199–202); Bleich and Pekkanen (2013), and Lynch (2013).

<sup>45</sup> Interview, LM-6, August 30, 2012. Wood (2006), for example, when conducting research on El Salvador's civil war, sought to interview people from organizations connected with the government *and* those aligned with the opposition.

<sup>46</sup> See Bleich and Pekkanen (2013, 87) on the weaknesses of snowball sampling and strategies to mitigate them.

<sup>47</sup> In any of these cases, one might employ a stratified sampling strategy (see, Rivera *et al.* 2002).

writing about and actors involved in the dynamics of interest,<sup>48</sup> and then divide them into categories (government officials, NGO activists, and party leaders, for instance). The categories might be used to organize a series of stratified focus groups. Alternatively, scholars could further divide those categories into types, and then choose a sample of people within each type to interview.<sup>49</sup> Scholars hoping to interview those whose voices are missing or have been systematically excluded naturally face particular respondent-selection challenges. Nonetheless, the more they learn about the research context and the phenomena under study, the easier it will become for those scholars to identify voices they *ought* to be hearing, and are not.

By contrast (and referring back to the third scenario above), the researcher may wish to make claims about a broader group by identifying the opinions, views, or behaviors of a subset of that group through interviews, and then generalizing to the broader group, or may wish to *test* causal explanations. When this is the objective, the ideal technique for choosing interview respondents may be to draw a random sample, as doing so enhances external validity. Drawing a random sample entails obtaining a list of the population of interest (or creating one if existing lists are unavailable, outdated, or systematically biased) in order to create a sampling frame. Given their time-consuming nature, researchers using in-depth interviews, focus groups, and oral histories are less likely to interview a sizable number of people, and thus less likely to need or wish to draw a random sample. As such, we leave discussion of this sampling strategy to Chapter 8, which examines survey techniques.

Each sampling or selection technique has strengths and weaknesses. To mention just a few, as noted above, while random sampling augments external validity, it may not be desirable for scholars who need to find out particular pieces of information that only a handful of actors might have. By the same token, while non-probability sampling does not mean non-careful, non-thoughtful, or non-justified sampling, there is always a risk that scholars who use such sampling techniques will accidentally exclude an important perspective or miss an important narrative strand.<sup>50</sup> If one engages in

<sup>48</sup> Scholars might even think of this selection process in terms of the dependent and independent variables in their study, and seek to identify individuals who can offer a range of perspectives on what the value of each variable is; doing so implies the need for strong theory in order to identify those variables (Cammett 2013, 136–137).

<sup>49</sup> Bleich and Pekkanen (2013, 90–91) offer a full description of this respondent-selection strategy.

<sup>50</sup> For instance, a scholar doing research in a remote village in Kenya, rather than accepting a sub-chief's offer to nominate interview respondents, found a list of farms and took a random sample of farmers to interview. The choice proved wise. While the researcher had been assured by the sub-chief that he could find everyone in his jurisdiction, the sub-chief did not know at least two-thirds of the farmers in the researcher's random sample (personal communication, August 23, 2010).

snowball sampling, for instance, the implicit level of trust between the interviewer and the respondent may be raised, but particular categories and types of respondents may lie completely off the snowball's track. Scholars should carefully consider these strengths and weaknesses before making a final decision about how they will select respondents, as their sampling choices have significant implications for the information they gather, and the role interview data can play in analysis and theory building (Lynch 2013, 38–44). Moreover, no matter which technique they adopt, scholars should systematically document each sampling and selection decision they make – describing their choices, highlighting the tradeoffs they faced, and justifying their final decisions.

We recommend that scholars seek to learn as much as they can about their chosen respondents before beginning to interview – for instance, investigating their positions on the phenomenon of interest, and their situation in their immediate milieu and the broader political context. Doing so helps scholars to decide how to contact desired respondents; to recognize the (potentially multiple) ways in which those respondents are connected to their study; and to pinpoint the types of questions that it will be most useful to ask (i.e., on what events, topics, or controversies the subject can shed light). Scholars conducting oral histories often consolidate this information in a biographical background sheet. Learning about the interview participants can also help the researcher to “market” herself and her project appropriately to different respondents (academics vs. policymakers vs. activists, for instance), identifying what to highlight and what to de-emphasize with each.<sup>51</sup> If there is an opportunity, subtly conveying her familiarity with a respondent (and his accomplishments) can demonstrate the researcher’s interest and seriousness of purpose.<sup>52</sup> This background knowledge can also enable focus group facilitators to pick up on subtle cues from participants and foresee potential roadblocks to the group’s conversation.

### **Constructing interview protocols**

The questions scholars ask and how they ask them influence the quality of their interview data and the way they can be employed in an analysis.<sup>53</sup>

<sup>51</sup> See Dexter (1970, 37–38) for an interesting discussion of pitching one’s project.

<sup>52</sup> Although see Rivera *et al.* (2002, 685) on how demonstrating knowledge of respondents may raise questions about anonymity.

<sup>53</sup> Through this subsection, we use the term “interview protocol” for ease of exposition. We neither have specifically in mind, nor mean to suggest that our thoughts only apply to, a highly structured questionnaire; the general principles we offer should be equally applicable to many types of interview guides.

Accordingly, many interviewing manuals offer rules and guidelines about protocol content, wording, tone, and question order (Weiss 1994; Leech 2002a, 667–668; Rubin and Rubin 2004, 146–164, 176–200; Wang and Yan 2012). Below we discuss how factors related to the field context shape protocol construction – for instance, concerning what language is used,<sup>54</sup> what questions are posed and in what order, and how they are phrased.

For more positivist scholars, writing an interview protocol that will generate useful data requires having a clear idea of what information they wish to gain through the interview and how it will be used in the analysis.<sup>55</sup> For these researchers, the research design, field research design, and interview protocol design processes can and should inform each other: the third is a way to translate the first two into action. Writing interview protocols forces a scholar to consider carefully what she will be listening for in interviews, what she thinks she might hear, and how they are linked to her project. Interview questions often connect to some aspect of a study's analytic architecture – they measure a variable, explore the nature of a concept, try to get at a causal process, or aid in case selection, for example. Accordingly, writing them helps researchers clarify their thinking on those aspects of their projects. If a scholar is developing a data-collection plan (discussed in Chapter 3) as he maps out and executes his research, he ought to be able to map each question, and the information he hopes it will generate, back to some aspect of that plan.

We suggest beginning to draft interview protocols well in advance of interviewing, for several reasons. Doing so provides scholars time to revise their protocols thoroughly and solicit local input. Writing and becoming familiar with their interview protocols in advance frees scholars to focus more on their respondents and the information they convey in the interviews themselves. Intense familiarity with the question guide is particularly crucial for focus group moderators, who must pay close attention to the words and body language for up to a dozen people simultaneously. Further, the more familiar a scholar is with his protocol, the more easily and quickly he can

<sup>54</sup> If the language or languages spoken in a scholar's field sites is or are not his native tongue, several factors might affect his decision about what language to use in an interview, including his skills in the local language(s), and what effect interviewing in his native tongue will have on the availability of respondents, on the conduct of the interview, and on the quality of the resulting data. The answers to these questions may vary across contexts, social groups, and time. Fujii (2013) discusses interviewing in a language one does not speak and using interpreters.

<sup>55</sup> Of course, we continue to advocate flexible discipline: scholars should not tie any interview to any particular analytic goal in their minds *so closely* that they miss relevant interview insights *not* tied to that goal.

identify essential questions not yet asked should a respondent or focus group member suddenly announce that she only has five more minutes to spend with the researcher.

Writing effective interview questions on the basis of one's research design or data-collection plan is difficult. How can scholars translate their key concepts and variables into concrete indicators, and then work those indicators into questions that respondents can answer in a way that helps scholars accomplish their analytic goals? We offer a few pointers. First, interview questions should be theoretically motivated but expressed in colloquial prose free of jargon so that interviewees can readily understand them. For instance, if the researcher is interested in whether a particular association is involved in broad or dense associational networks, he likely cannot ask an association leader, "Do you consider yourself well integrated in an associational network?" or follow up with, "Would you characterize that network as broad, or dense?" However, he could ask "How many other associations do you work with?"; "What kind of work do you do with them?"; "How often do you work with them?" The researcher can then evaluate the depth and breadth of associational networks based on information garnered through these straightforward and easily intelligible questions.

It is also important to identify colloquial language that is appropriate to the setting and the interlocutor or focus group. For instance, in a country recently transitioned from military rule, a researcher might pose a question about human rights one way when interviewing a human rights activist, and another way when speaking with a military officer. Moreover, scholars should keep in mind that their respondents' frames of reference for particular issues can be very different from their own. For instance, one researcher we interviewed noted that the Peruvian street vendors she had interviewed generally thought of income in three-hour increments (the time they might spend at any one post) rather than in terms of an annual salary.<sup>56</sup> Scholars can offer context-sensitive stimuli without changing the meaning or substance of their questions, and posing questions in terms respondents understand and do not reject will generate better data (see Locke and Thelen 1995).

As scholars develop their interview protocols, it can be useful to pre-test them. Pre-tests may occur in different ways depending upon one's research goals. Scholars with clear hypotheses and highly structured protocols may be able to pre-test them much as one would pre-test a survey questionnaire, using pre-interview trials and/or open-ended discussions. This sort of more

<sup>56</sup> Interview, DK-1, July 20, 2012.

structured pre-testing can be particularly important in connection with focus group interviewing, as moderating focus groups requires such intensive preparation and training: pre-testing allows for the evaluation of the perceived meaning and intelligibility of the questions that will guide discussion, and provides the moderator opportunities to practice and strengthen her facilitation skills. Ideally, pre-tests are done on a sample of the target population, but they need not be. Many wrinkles, from question wording and order, to phrasing and language, can be ironed out with just a few runs through interview questions with a native speaker of the language in which the protocol is written (particularly if it is not the researcher's native tongue) or someone familiar with the local context.

For others, pre-testing will be understood more loosely, as a way to examine whether questions elicit the kinds of answers a researcher anticipates and avoid provoking negative responses. For instance, one might "float" particular concepts or expressions to see how individuals think about them and to inform how one might ask questions about them. Scholars can often benefit from honing the appropriateness of the language used in their questions by seeking the input of local researchers. Even if these individuals have not done interviewing themselves, they can often evaluate whether the questions a scholar plans to ask are clear and likely to prompt respondents to provide the desired information. If a scholar has contracted research assistants, they may also be able to help.

Of course, most researchers' interview protocols change over the course of their fieldwork as they learn more about the field site, their topic, and their respondents, and apply what they are learning to their original categories or hypotheses. Even with careful pre-testing, scholars often realize that questions need to be added, modified in order to elicit the sought-after information, or abandoned completely because they just "don't work" or are attached to lines of inquiry that have proven to lead nowhere.<sup>57</sup> As we have emphasized previously, researchers should not interpret the iterative nature of fieldwork as a reason to dispense with advance planning, and carefully preparing their interview questions is no exception. Further, as the principle of flexible discipline would suggest, scholars should carefully evaluate the implications of any changes they make to their interview strategies and protocols before they make them – and fully document those choices, their justifications, and their analytic ramifications – so they can be as transparent as possible in their eventual write-ups. And, of course, as

<sup>57</sup> Interview, LM-15, September 10, 2012.

scholars develop new hypotheses and de-emphasize others, they should continue to search for evidence, through interviews and other means, that both support and refute their old and new hypotheses, as well as rival ones.

### Gaining access to interview respondents

Once interviewees have been chosen, the next challenge is to gain access to them (Goldstein 2002).<sup>58</sup> Encouragingly, respondents are sometimes quite eager to speak with researchers – particularly ones who are interested in what they do and seem understanding, non-judgmental, and trustworthy.<sup>59</sup> Having a connection to a researcher may be useful for some respondents, and others may be very keen to tell their story or their side of a controversy, to unburden themselves, or to expound on their accomplishments or on a topic about which they feel expert (Dexter 1970, 5, 36). Yet most scholars experience some challenges in accessing the people whom they would like to interview in the field. Indeed, for many scholars, gaining access is one of the most frustrating (albeit ultimately rewarding) aspects of field research.

Many factors can complicate accessing potential respondents for interviews.<sup>60</sup> Scholars may be unable to locate those whom they wish to interview. Alternatively, potential respondents may feel uncomfortable being interviewed (or talking about the specific topic at hand), may seek to signal importance by being aloof, or may react negatively to the researcher.<sup>61</sup> Reaching desired interview subjects may be especially difficult in authoritarian and highly controlled or dangerous environments where additional obstacles can arise even once the interview is granted.<sup>62</sup> Not being able to

<sup>58</sup> For additional guidance on recruiting respondents, see Dexter (1970, 46–50); Weiss (1994, 33–37); Rubin and Rubin (2004, 93–97); and Beckmann and Hall (2013, 199–202); scholars might also consider recruiting respondents through social networking sites.

<sup>59</sup> Moreover, reluctant respondents can sometimes become enthusiastic ones; one of our interview respondents recounted that simply re-framing her interviews in Chile as *consultas* (more like “questions”) rather than *entrevistas* (“interviews”) gave the sense that respondents would be teaching, made the encounter sound safer, and encouraged potential respondents to grant interviews (interview, DK-17, August 24, 2012).

<sup>60</sup> Note that many of the access challenges we discuss here are analogous to those considered in Chapter 5.

<sup>61</sup> The PS symposium edited by Ortbals and Rincker (2009) offers a range of first-hand accounts of how identity affects access to interview respondents; for additional guidance on such points see Dexter (1970, 46–50); Weiss (1994, 33–37); Rubin and Rubin (2004, 93–97); and Beckmann and Hall (2013, 199–202).

<sup>62</sup> Official “minders” may be assigned to accompany and observe respondents during interviews, or access may be state-regulated. One interview respondent recounted that she had to solicit permission from neighborhood officials in order to interview ordinary people in China (interview, DK-18, August 24, 2012).

reach desired respondents can sometimes have analytic implications. For instance, if a scholar is seeking to develop claims that are generalizable to a larger group on the basis of interviewing a sample of that group, and if many individuals in the sample who are systematically different from the rest of the sample in some way cannot be interviewed (and if they are not systematically replaced), non-response bias can result (see Goldstein 2002, 669–670, for a good discussion).

With regard to strategies for circumventing obstacles, we first suggest that scholars assess precisely how necessary it is to speak with the individuals who have proven difficult to nail down for an interview. How will the study's results be affected if those people cannot be interviewed? Can the data that would be gained from interviewing them be obtained in some other way? How will not reaching them affect the quality and quantity of, and a scholar's confidence in, her evidence? How will it affect her ability to build theoretical arguments? Scholars should also consider what they can learn – with regard to both their project and future contacting efforts – from who turns down their interview requests, and from the way in which and frequency with which they do so (Dexter 1970, 37).

Should scholars decide to pursue elusive respondents, identifying a connection to them – an associate, relative, friend, colleague, or employee, for instance – can prove invaluable for facilitating access.<sup>63</sup> Scholars might also establish an institutional affiliation, or more informally ask for advice or introduction from scholars working on similar topics in the field site (Solinger 2006, 157). Likewise, scholars should remain attentive to opportunities to enter the networks in which target respondents operate (e.g., they might seek to attend the types of events they attend, or get on mailing lists to which they subscribe). Reinhhardt (2009, 296), for example, recounting her field research experiences in Brazil, noted that “people who had no interest in returning my phone calls, upon meeting me at a party, would instantly invite me to their office the next day” – perhaps, she surmises, because she passed some sort of “test of . . . sincerity and commitment to the person who had invited me.” Connections made using these techniques can very quickly generate a response to an unanswered email or phone call.<sup>64</sup>

<sup>63</sup> Interview, LM-5, August 27, 2012.

<sup>64</sup> See also Cammett (2013) on proxy interviewing (e.g., contracting others to do interviewing), especially her discussion of training and managing proxies; scholars considering this option whose research takes place in a country that is not their own should determine whether their research grant precludes paying foreign nationals.

While it may often be easier to arrange interviews with “ordinary people” than with elites, this is far from a universal truth. Indeed, a range of additional challenges may arise when scholars seek to access non-elites for interviews. Ordinary people may be completely consumed with their complicated lives, may have little motivation to convey personal information to researchers, and may be more nervous or suspicious about, and less cognisant of the value of, doing so (Cammett 2013, 128). As such, personal connections can be just as important to facilitate the trust necessary to secure interviews with non-elites. For example, one of our interview respondents needed an insider to “walk him into and around,” and introduce him to people in, the several shantytowns in Argentina where he wished to study patron-client relationships. Without that “human imprimatur,” he suggested, it would have been unlikely that anyone would have spoken with him (and quite likely he would have been ushered out quickly).<sup>65</sup> Likewise, in Yashar’s (2005) work on indigenous politics in the Andes, a small group of indigenous activists in each country she studied invited her to their homes, to conferences, and to workshops, helping to open doors to other indigenous leaders and activists.

In short, the more networking scholars do, the more likely they are to identify valuable connections. While connections sometimes appear out of the blue, often researchers have to work to create, cultivate, and maintain them – and the time they anticipate spending doing so should be factored into their plans.<sup>66</sup> Employing a variety of intermediaries and access points can prevent scholars from becoming trapped in a subset of the larger universe of potential respondents.<sup>67</sup> One of our interviewees realized that she needed to “get off the beaten path” (move beyond the NGO world) to access a more diverse group of resident association leaders in Rio de Janeiro, and did so by “striking up conversation with whomever [she] could” – people in the stores in which she shopped, on the beach, etc. – in order to develop broader networks.<sup>68</sup> Scholars should also keep in mind that the way in which they contact respondents creates an impression and may lead interviewees to draw particular conclusions about the researcher that can affect the information they will offer: like everyone else, scholars are often judged by their associations.<sup>69</sup> One political scientist we interviewed, for instance, was

<sup>65</sup> Interview, DK-16, August 21, 2012.

<sup>66</sup> Sometimes respondents in one research project become connections for the next; see Solinger (2006) on the usefulness of maintaining contacts.

<sup>67</sup> Interviews, DK-13, August 8, 2012; DK-19, August 27, 2012.

<sup>68</sup> Interview, DK-13, August 8, 2012.

<sup>69</sup> Interviews, DK-11, August 7, 2012; DK-12, August 8, 2012.

strongly advised *not* to affiliate with any particular research institution as doing so would have pigeonholed her – perhaps opening some doors wider but firmly closing others.<sup>70</sup>

Researchers who are interested in a particular organization, group, or community of people should likewise carefully consider how to “enter” – what cultural etiquette to follow, and whether to enter from the top, the bottom, or via a particular gatekeeper. In some government bureaucracies or communities, for instance, it is critical to start at the top of the hierarchy and obtain official approval to interview respondents at each lower level.<sup>71</sup> In other cases, political hierarchies may be less rigid or less transparent but researchers will nonetheless need a particular person or people to aid them in gaining entrée. And sometimes starting at the “bottom” and proving oneself to those in lower echelons is critical to eventually gaining access to someone at the pinnacle. Moreover, sometimes lower-level bureaucrats, deputies, or staff know more about the workings of an organization – and are more forthcoming – than those at the top. As with many other points we have discussed, the chain of command, and what type of person could most effectively serve as a key facilitator or broker, will differ significantly from context to context. This again highlights the importance of scholars investing the time necessary to understand and engage with their field sites.

### **Conducting interviews: interviewer effects, rapport, and the ethics of interviewing**

The emerging political science literature and the vast sociology, anthropology, and psychology canons on interviewing offer a wealth of useful information about structuring and conducting interviews, addressing topics such as deciding upon location, considering whether to send interview questions in advance, evaluating the need to offer gifts to respondents, introducing oneself and one’s work, pacing, re-capturing a wandering interlocutor, probing, asking follow-up questions, wrapping up the interview (or reacting to its early termination by a respondent), requesting additional contacts, thanking respondents, and more.<sup>72</sup> We limit our comments here to three

<sup>70</sup> Interview, DK-17, August 24, 2012.      <sup>71</sup> Interview, LM-13, September 7, 2012.

<sup>72</sup> See, e.g., Dexter (1970, 50–54, 58–70), Weiss (1994, 61–119), Aberbach and Rockman (2002), Berry (2002), Leech (2002a), Rivera *et al.* (2002), Rubin and Rubin (2004, 135–146, 164–176), Kvale and Brinkmann (2009, 123–141), Herzog (2012); in political science, see Beckmann and Hall (2013, 288–296) and Leech *et al.* (2013). On oral history interviewing, see Dunaway and Baum (1996), Perks and Thomson (1998), Thompson (2000), Yow (2005) and a host of online resources available through various oral history projects and centers (see The Oral History Association’s website,

ways in which the field context influences the conduct and products of interviews: interviewer effects and positionality, rapport, and ethics.<sup>73</sup>

Many scholars believe that a researcher's self-presentation and personality, and a respondent's perception of the interviewer's identity and personal traits (e.g., her gender, ethnicity, marital status, educational level, socio-economic status, and nationality) – in interaction with the research context – shape the interpersonal dynamic of an interview and thus the data collected through it (Dowling 2000; Ortbals and Rincker 2009, 287).<sup>74</sup> Moreover, as discussed in Chapter 4, the effect that particular aspects of a researcher's identity can have on interviewing may be neither predictable nor stable for any one project, let alone between projects (Becker, Bonnzaier, and Owen 2005), and those traits may gain and lose relevance across interviews (Chacko 2004). Several researchers we interviewed remarked how being married, or having children, elevated (and diminished) their status and facilitated (and complicated) mutual respect in their interviews. Others expressed surprise at how they were welcomed or rejected as an "insider" versus an "outsider" in their field site over time. To offer one vivid example of this last dynamic, a scholar of Serbian ethnicity doing research on human rights in the Balkans was harassed and threatened after she criticized Kosovo's declaration of independence during an appearance on CNN. Through a textual analysis of the hate mail and online postings she received, she realized that her correspondents' hostility was rooted in a primordial view of ethnicity that led to her Serbian ethnicity "trumping" other aspects of her identity: "Try as I might to de-ethnify myself, to cloak

[www.oralhistory.org](http://www.oralhistory.org)). On focus group interviewing, see Greenbaum (2000), Krueger and Casey (2009), and Stewart and Shamdasani (2014).

<sup>73</sup> Chapter 4 offers additional information on these topics. See also Dexter (1970, 32–34, 115–128) on a variety of interviewing relationships and a "transactional theory" of interviewing; Mishler (1991, 52–65, 117–135) on the "joint construction of meaning" in interviews and the distribution of power between interviewer and respondent; Woliver (2002) on ethical dilemmas in interviewing; Rubin and Rubin (2004, 79–128) on building "conversational partnerships," "responsive interviewing," ethics, and the stages of interviewing; Warren (2012) on interviewing as a social interaction; Lillrank (2012) on "Managing the interviewer self"; and Cammett (2013, 126–128) on positionality in interview-based research. Kaiser (2012), Marzano (2012), and Miller-Day (2012) offer very good discussions of informed consent, protecting confidentiality, and IRBs, and Kvale and Brinkmann (2009, 61–79) address the central moral and ethical concerns in interviewing.

<sup>74</sup> These issues, captured by some via the term "positionality" and by others via the term "interviewer effects," can be relevant no matter whether the researcher's power is by some measure weaker, on a par with, or greater than that of the respondent. Of course, these effects do not impinge only *during* interviews; a scholar's identity (and her biases) shape the questions she asks and how she asks them, as well as how she perceives and interprets the answers respondents provide (Dexter 1970, 125; Townsend-Bell 2009, 311).

myself in the robes of an academic scholar, ethnicity comes back to define both me and my work" (Subotic 2010).

While there is some consensus among political scientists that these effects occur, positivists and interpretivists differ in their views on how such effects shape the interview exchange and influence the data derived from it, and what can or should be done about it. Scholars engaging in reflexivity explicitly acknowledge the impact identity has on the research experience and weave accounts thereof into their findings and written products.<sup>75</sup> More positivist scholars may also actively seek to identify and estimate how their identity affects data collection – but consider it a source of bias or measurement error (Mosley 2013b, 12–14). One senior scholar described "yearning for" objectivity: "Even though I'm not sure it's possible. I keep working to get rid of everything that is contaminating and all of the biases to get closer to some idea of the truth."<sup>76</sup>

Another aspect of the interaction between the researcher and the respondent that affects interviews and the information gathered from them is whether and how researchers seek to build rapport with and gain the trust of respondents, and how successful they are. After all, negotiating access does not end once a researcher has secured an interview with a desired individual or group. Access to the *information* the scholar is after continues to be negotiated during interviews. The better the researcher understands how respondents see situations (through building rapport), the better he will be able to word or target his questions, and interpret the answers respondents give (Dexter 1970, 119–120). Focus group interviews, for instance, usually begin only after some informal mingling and then a round of introductions so participants can establish rapport with the moderator and each other. Solinger suggests researchers try to establish "mutual authority" in the interview context (2006, 163) – an understanding that both interviewer and interviewee are driving the conversation. A scholar who interviewed actors involved with political transition in South Africa had the palpable feeling that potential respondents needed to understand her feelings before opening up to her – they wanted "some small sign that I 'got them.'"<sup>77</sup>

<sup>75</sup> Finlay (2012) offers an enlightening discussion of reflexivity; see also the extensive symposium on fieldwork, identity, and intersectionality edited by Ortbals and Rincker (2009).

<sup>76</sup> Interview, LM-18, September 14, 2012.

<sup>77</sup> Interview, DK-10, August 6, 2012. Leech (2002a, 665–666) suggests several additional strategies for building rapport, and Chakravarty (2012) offers a detailed, and potentially controversial, discussion of building trust.

Of course, researchers should not be so concerned with establishing rapport that they lose objectivity, or the ability to keep interviews on-topic (Rivera *et al.* 2002, 685). One researcher working in China explained, “I try to be empathetic, not sympathetic . . . I try to understand what the world is like through their eyes but try to avoid representing them.”<sup>78</sup> Grele describes how the friendly rapport established doing fieldwork can lead to biased questions, answers, and interpretations in oral history interviews:

On both sides of the microphone, to ease the social situation, to maintain empathy and rapport, we avoid the hard questions and the unsettling answers . . . History without biases and passions is probably impossible and if attainable would be as dull as dishwater. But in doing our fieldwork we must overcome the natural tendencies of social intercourse and remember that we are historians and we are interested in the fullest exposition of the passions of the past, not in gathering material which is acceptable to the present. (1978, 41)

Moreover, particularly when scholars are interviewing people on different sides of an issue (who may nonetheless know each other), carefully balancing empathy and neutrality helps them to avoid poisoning the waters for future interactions.

Perceptions that individuals in the broader field context hold of researchers can also affect interviewing. In some field situations, for instance, being seen with a researcher could endanger interviewees. As Hsueh (2008) writes concerning her fieldwork in China, because she is a foreigner, her being detained by state security when researching sensitive topics would likely lead to no more than a session of detailed and intimidating questioning by a team of “good and bad cops.” For those whom she was interviewing, however, the consequences might be far worse. As we have insisted, researchers have an ethical responsibility to protect study participants from harm.<sup>79</sup> Fulfilling that responsibility entails thoroughly investigating the politics of the field context, coming to understand the contingent implications for participants of being involved in the project, and thoughtfully considering whether particular individuals should be excluded from the study, or particular questions left unasked.<sup>80</sup>

In many situations, of course, being associated with a researcher would not be dangerous or stigmatizing for respondents. On the contrary, sometimes

<sup>78</sup> Interview, LM-6, August 30, 2012.

<sup>79</sup> For additional discussion of several of the points concerning ethics made here, and others, see Chapter 4.

<sup>80</sup> Interview, LM-8, August 30, 2012.

respondents anticipate that such associations will bring benefits for them or their community, and those expectations must be dealt with ethically. As we discuss in more detail in Chapter 4, it is important to critically reflect on whether and how to give back to participants without aggravating inequalities or inducing envy. For researchers working in relatively impoverished field sites, respondent or community expectations and demands can represent a significant emotional drain.

Carefully adhering to the process dictated by their university IRBs can help researchers to navigate these sorts of ethical challenges.<sup>81</sup> Yet ethical interviewing may entail a more encompassing ethical commitment (MacLean 2013). For instance, before arriving for an interview, researchers should consider whether they will need to gain consent from the respondent alone, or whether other individuals who are superior in an organization – or other groups and communities – may need to provide consent and approval. MacLean found that for a study of American Indian representation in health policy, she needed to obtain nearly fifteen review-board approvals, including from her university IRB, federal and regional IRBs for the Indian Health Service agency of the US government, and formal and informal tribal IRBs for participating tribal nations. To offer another example, for focus groups, researchers often ask all participants to sign a group agreement that she also signs (rather than simply signing an individual consent form as with an individual interview or oral history).

Before an interview begins, it is likewise incumbent upon the researcher to find an appropriate location for it. This is important both because where an interview is conducted (and, accordingly, how comfortable the respondent is) affects the data it generates, but also because finding a private location is often key to assuming the ethical responsibility for ensuring confidentiality if such guarantees were made.<sup>82</sup> Even when a quiet spot is found, curious children, spouses, friends, or colleagues may attempt to drop in, interject, or just quietly listen or observe from afar. Researchers should anticipate these interruptions and develop a culturally appropriate script for turning interlopers away. This can be especially delicate when the subject (or the researcher) is more junior in age, social or political status than the intruder

<sup>81</sup> See Brooks (2013) for an illuminating overview of human subjects protection.

<sup>82</sup> To clarify, by “private” we do not necessarily mean “isolated”; a restaurant with few patrons – or indeed a busy restaurant in which no one else could possibly hear the conversation – could be considered “private.”

and particularly difficult with a focus group that appears to be an informal, lively discussion among peers.

Further, the process of soliciting informed consent is more than a bureaucratic hurdle to be surmounted by rushing respondents into signing consent forms or offering oral consent. Researchers should strive to make the potential risks *and* benefits of participation as concrete and transparent as possible for potential human subjects. Respondents should know that they can refuse to answer certain questions and may stop the interview at any point. In a focus group, this information is usually conveyed by the moderator at the beginning in a brief discussion of the ground rules; she might also explain that the focus group's objective is to reveal multiple perspectives and not to impose a consensus. Also, the researcher should be aware of, and remain attuned throughout the interview to, cues that suggest discomfort or refusal, even when participants continue to answer questions.<sup>83</sup> This requires particular sensitivity and extra effort in a focus group, where members may feel social pressure to continue participating after agreeing to do so. Moreover, should a researcher believe (or come to believe) that certain information an interviewee conveyed is too sensitive or controversial to be made public or attributed to him or her, it is the researcher's responsibility to protect the respondent, perhaps keeping the information confidential even when the respondent did not request confidentiality.

When guarantees of confidentiality were extended for some or all of the information conveyed in an interview, it is incumbent upon researchers to make the necessary arrangements to maintain those guarantees. Given the richness of data gained through interviewing and the amount of explicit and implicit identifying information found in interview notes, tapes, and transcripts, all of these materials need to be stored in a way that will ensure confidentiality (as required by IRB guidelines). Interview transcripts should include a code rather than the respondent's name, and codebooks with identifying information should be carried and stored safely and separately from the coded interview notes or transcripts. Keeping interview data private and confidential when transporting them across long distances and over borders can be especially important and challenging if data are sensitive – or become sensitive (and sought-after, for instance, by government entities) – due to changing political conditions in the field site.<sup>84</sup>

<sup>83</sup> For an eye-opening discussion of the ethical responsibilities entailed in asking interview respondents to relive traumatic experiences, see Wood (2006).

<sup>84</sup> Interview, DK-12, August 8, 2012.

Interviewing in an ethical fashion, in short, means more than simply mouthing one's IRB protocol. Indeed, as we suggest by including "ethical commitment" as one of fieldwork's six guiding principles, we believe ethics are a "way of being" in the field, entailing researchers consistently following the spirit of their pledge to protect those whom they involve in their work. With respect to many of the issues discussed in this and the previous subsection, asking other researchers who have worked in the context or similar contexts about how they faced challenges with ethical implications, and what solutions they devised, can be very helpful.

### Capturing information from interviews

In deciding how to capture information from their interviews, researchers face additional important choices: they might audio- or video-record, take notes, do both, or do neither. Several factors may affect choices about capturing interviews (see also Dexter 1970, 54–57; Weiss 1994, 53–57). Sometimes logistical factors dominate capture choices. For instance, scholars may wish to audio- or video-record focus group sessions, as it can be difficult to coordinate the group, employ a coding scheme to track individual participants' contributions, and take notes simultaneously. Moreover, interactions in focus groups can be lively, animated, fast-paced, and nearly impossible to capture with any accuracy without recording.

More often, however, scholars will have a choice about how to capture the content of an interview. When they do, they should consider what information they would like to take away from their interviews, how much information, and in what form. Recording provides a full and faithful record of everything said – a very rich data source that can be used to confirm or clarify points later. By the same token, the amount of data gained can be overwhelming, and can complicate the process of drawing on interviews in later phases of analysis and writing. Anticipating how they will analyze and use information gained from interviews in their study can inform scholars' data-capture choices. If it is important to have long, exact quotes, or if content analysis will be employed to analyze interview data (suggesting the need for a verbatim text), recording may be useful. Of course, even if interviews are recorded, it can be prudent to write down interviewees' responses simultaneously so interview data exist in more than one form. Scholars who take notes while conducting less-structured interviews can also jot down additional questions that occur to them as the interview progresses.

Regardless of how scholars decide to capture respondents' (and their own) spoken words, additional aspects of an interview, and the thoughts and observations that occur to scholars as they conduct it, can also be analytically significant and important to track. For example, information about the context or venue of the interview (and why it took place there); who else was present and for how long; the length of the interview (and why it was as long or short as it was); and researchers' reactions to particular interview responses or to the interviewee's general demeanor, can all be important. Likewise, what was said "off the record," "not for attribution," or "on background" need to be noted.<sup>85</sup> In focus groups, observations about intra-group interactions are critical to record. These "interview metadata" can enrich the value of an interview, serving as a useful memory prompt, contextualizing the information conveyed, and enhancing scholars' ability to interpret it and assess its evidentiary value. Scholars might insert such ideas and reactions right into their interview notes, taking care to distinguish between what the respondent said and what their reactions were (perhaps using a double-entry system), or might create an "interviewer observations sheet" to capture these metadata.<sup>86</sup>

The way in which a scholar decides to capture interview content can affect the data he collects. Some respondents (for instance, those who speak to the media frequently) may feel entirely comfortable being recorded, and may even prefer being recorded. Others may withhold information if a tape recorder is running, or become unnerved by being recorded while interviewers simultaneously write out their responses.<sup>87</sup> While it will not always be possible to determine respondents' preferences, as noted previously, they will often – intentionally or inadvertently – signal their comfort level, perhaps as the researcher offers IRB-mandated statements and inquires about recording and confidentiality. Researchers should be attentive to such cues.

Scholars who take notes during interviews should type them up as soon after the interview as field circumstances permit.<sup>88</sup> Doing so optimizes the quality of the data gathered from the interview. If handwritten notes from an interview cannot be typed up quickly, cleaning them up quickly – filling in

<sup>85</sup> See Goldstein (2002, 671) for a discussion of the distinctions among these.

<sup>86</sup> This practice is more commonly associated with survey research than with in-depth interviewing.

<sup>87</sup> While researchers must *always* reveal their intent to record, using pens that are simultaneously audio recorders can deemphasize the presence of the recording device.

<sup>88</sup> Another option is to voice-record details about the interview immediately afterwards; these recordings can be run through a voice recognition program, producing a digital text version of interview notes and impressions. Products such as LiveScribe quickly convert handwritten notes on paper into text on an iPhone or iPad, facilitating organization, tagging, and searching.

unfinished words and half-written sentences – will facilitate typing up the notes later. If interviews are recorded, scholars will need to decide whether to transcribe them.<sup>89</sup> Doing so can be time-consuming or expensive – or both if researchers spend time recruiting<sup>90</sup> and supervising transcribers.<sup>91</sup> A less time-consuming alternative might be to construct a “table of contents” summary of an interview recording by listening to the recording and making a simple notation, every minute or two, concerning the topic of discussion, and adding notes highlighting particularly interesting quotes, or other points they may want to be able to easily identify later. Another option – if good notes were taken during the interview – is to set the recordings aside until the writing stage, when they may prove useful to draw a specific quote or review exactly how a respondent answered a question. Points for which it might be useful to have an exact quote later can be flagged on interview notes.

### **Assessing and documenting the evidentiary value of interview data**

As we suggested in the previous chapter, no source of information can be assumed to provide “the truth.” Summarizing the challenges of interviewing, one scholar stated, “You have to accept that people are going to be telling you their version of the truth.”<sup>92</sup> Respondents are not obliged to be objective (Berry 2002, 680); indeed, researchers across the epistemological spectrum question whether they can be. Respondents can vary considerably with regard to their knowledge of the events and circumstances researchers query them about, and how clearly they recollect history. Moreover, as the literature on oral history emphasizes, while respondents are being asked to recall the past, they are being interviewed in the present, meaning that the history they recount is unequivocally shaped by their current context and situation. Subjects also differ in terms of how reflective, trustworthy, forthcoming, and given to hyperbole they are. One-on-one with a researcher, people may seek to depict events or dynamics in a certain way, aggrandize their role and downplay that of others (or vice versa), and leave out important information.

<sup>89</sup> Kvale and Brinkmann (2009, 177–187) offer a useful primer on transcribing.

<sup>90</sup> Scholars might contract graduate students recommended by faculty contacts, or hire a commercial transcription services provider, either at the field site or after returning to their home institution. Such services can be quite expensive in the United States.

<sup>91</sup> Transcription software is continually developing. Researchers should be sure to investigate the quality of the transcriptions such software produces in the language or languages in which they operate.

<sup>92</sup> Interview, LM-15, September 10, 2012.

Interviewees can lie, more or less blatantly.<sup>93</sup> And, as we noted, respondents are often influenced by their perceptions of the researcher: the same person interviewed about the same topic by two different scholars may offer different information.

These possibilities and limitations are among the strongest criticisms that skeptics level at research based on interview data. If respondents and interviews vary on these parameters – and if there is no clear metric to evaluate whether scholars have interviewed those who can provide the most accurate information concerning their question – what can interview research contribute to a discipline focused on objectivity and generalizability?<sup>94</sup> Interview research raises issues of validity (how effectively the “measuring instrument” – the interview protocol and the interviewer – captures the information it is designed to capture) and reliability (the potential consistency of repeated tests with the measuring instrument, i.e., whether the same answers would be produced if interviews were repeated) (Berry 2002, 679). We suggest four strategies researchers can use to address these potential weaknesses of interviews, and to evaluate the evidentiary value of the data collected.<sup>95</sup>

First, as mentioned above, researchers should seek to interview individuals from as many theoretically relevant categories, with as many differing perspectives or positions on the dynamics of interest, as possible. Doing so will help to attenuate non-response bias. If they are ultimately unable to interview many people, and can establish how those individuals are different from those with whom they spoke (something scholars conducting this sort of research may well be able to do), they might be able to estimate how the absence of those voices affects the data (Goldstein 2002, 672). In the text of their written products or in a methodological appendix (or both), scholars should outline how they sampled and contacted respondents; construct a log of whom they sought to interview, who was interviewed, who was contacted but failed to be interviewed and the reasons why, and how their failure to be interviewed might affect the conclusions that are drawn from the data; and identify the study’s response rate.

<sup>93</sup> For more on these issues, see Weiss (1994, 147–150); Fujii’s (2013) particularly useful, well-illustrated discussion; and Dexter (1970, 100–114).

<sup>94</sup> Several of our interview respondents suggested that such questions about interview research had been posed to them or about their work (e.g., interview, DK-9, August 21, 2012); see Kvale and Brinkmann (2009, 168–171, 293–298) for additional standard criticisms of interview research.

<sup>95</sup> Rubin and Rubin (2004, 71–77) and Kvale and Brinkmann (2009, 241–265, 298–315) discuss similar issues and offer additional advice.

Second, concerning those who *were* interviewed, and in line with our discussion of interview “metadata,” actively judging the credibility and reliability of respondents and of the information they provide is critical.<sup>96</sup> To do so, scholars might ask themselves a few sets of questions. Concerning the respondent, what degree of involvement in or exposure to the phenomena of interest did he have? Is he representative of the larger group or community to which he belongs?<sup>97</sup> What goals or ulterior motives might he have had in giving the interview? What biases did he convey?<sup>98</sup> Regarding the interviewing interaction more generally, scholars can think about how the interview context – and the interviewer – might have affected the interviewee and her responses. And concerning those responses, the researcher might consider at what points the respondent was hesitant or evasive (and why she might have been so), and assess how internally consistent, comprehensible, plausible, or self-serving her responses were. This information, perhaps in tandem with a comparative analysis of subjects and the information they provided, can also be included in scholars’ write-ups.

While not providing a fail-safe measure of the quality of the information conveyed through an interview, considering these aspects has two benefits. First, if a researcher can identify a respondent’s biases during an interview, she can try to counter them as the interview progresses (for instance, by indirectly asking the respondent to “critique his own case”) (Berry 2002, 680).<sup>99</sup> Second and more broadly, answering these sorts of questions allows the researcher to make informed judgments about how much distortion respondents might have introduced and in what direction,<sup>100</sup> and thus to evaluate the evidentiary value of the information provided. Doing so may direct the researcher to weigh some interviews, or certain points made in

<sup>96</sup> As Dexter (1970, 108) very aptly states, the key question is not whether respondents are telling the truth, but “What do the informant’s statements reveal about his feelings and perceptions and what inferences can be made from them about the actual environment or events he experienced?”

<sup>97</sup> Evaluating the representativeness of respondents who are very eager to be interviewed can be particularly important. Likewise, especially in oral history interviews, a handful of “survivors” (i.e., a small subset of those who experienced a traumatic event) may well be unrepresentative of the broader group (most of whom did *not* survive). For additional basic background on oral history interviewing, see [www/historymatters.gmu.edu/mse/oral/how.html](http://www/historymatters.gmu.edu/mse/oral/how.html).

<sup>98</sup> Much of this could be considered the “latent content” of an interview; see Lynch (2013, 36) on the distinction between overt and latent interview content.

<sup>99</sup> See also Wood (2006, 382) on whether to challenge interview respondents whom the researcher knows are lying.

<sup>100</sup> We can also think of these distortions as “measurement error,” or simply as having introduced “error” into the “model of meaning” the researcher is striving to construct (Dexter 1970, 110).

particular interviews, more than others, and allow him to generate estimates of how certain he is of findings based on his interview data.<sup>101</sup>

Third, no matter how central a figure is in a particular event, how authoritative a source seems to be, or how much confidence a scholar may have in her account, a single interview, focus group, or oral history is rarely sufficient to gain the data necessary to understand an event or support an argument. In the vast majority of cases, it is important that scholars acquire additional viewpoints and data through some combination of further interviewing and employing other data-collection techniques.<sup>102</sup> As emphasized throughout this book, engaging in triangulation is the ultimate antidote to doubts about data.

Finally, and to reinforce a point we have made several times, scholars should include as much information about the interviewing process as they can in their write-ups. In addition to the specific points highlighted above, they should offer a general description of how they conducted and captured interviews, document the parameters and resolution of all methodological problems encountered while interviewing (Berry 2002, 679), and perhaps provide the interview protocol. Bleich and Pekkanen (2013) offer a helpful discussion of these issues, and provide templates for creating what they call an “Interview Methods Appendix” and “Interview Methods Tables.” Making sure that this information becomes part of the research record – which will be easiest for scholars who begin writing about these issues *while in the field* – ensures transparency,<sup>103</sup> giving readers the tools they need to evaluate interview data, the process through which they were collected, and the conclusions drawn from them.<sup>104</sup>

---

## Conclusion

---

Conducting interviews enhances any researcher’s understanding of the substantive problem he is studying, key concepts, and causal processes. Choices

<sup>101</sup> On weighing interview data proportionate to confidence in their validity, see Berry (2002, 680) and Allina-Pisano (2009).

<sup>102</sup> Grele (1978, 30) urges historians to evaluate oral history sources critically and implicitly advocates triangulation when he declares, “it is not the case that all studies of the past done in the past are in error. Nor is it true that oral histories of the dispossessed are *ipso facto* free of the biases of the larger culture.”

<sup>103</sup> Golden (1995) discusses how transparency and confidentiality might conflict and argues that qualitative scholars should aim for replication-enabling transparency nonetheless; Mosley (2013a, 20–26) offers additional cautionary notes.

<sup>104</sup> See also Aberbach *et al.* (1975, 9–12) for ideas about what data about interviews to report and how to report them.

concerning whether, when, and how to engage in interviewing, and which techniques to use, are driven by a combination of factors – the researcher, the nature of the research question, research design, and field circumstances. Because each interview is a unique encounter in a particular time and place, no book can offer infallible solutions to every potential challenge researchers might face when conducting in-depth, oral history, and focus group interviews in the field. Instead, we emphasize the types of issues to consider when addressing these challenges. Scholars will be more successful at landing and conducting interviews in the field if they are clear about their analytic goals; familiar with interviewing methodology; understand the language, history, culture, and economy of the interviewing context; and actively engage with the field site. Preliminary field-site visits can provide knowledge that can help scholars to refine the research question, select cases, and develop the trust necessary to conduct interviews.

We hope we have made clear that the most important task for those conducting interviews in the field is listening. Scholars who become wrapped up in the data chase sometimes lose sight of the fact that their ultimate goals, as field researchers, are to listen, absorb, and learn. Relatedly, it is essential that the subject leave the interview feeling that he or she has been treated with respect. As one scholar eloquently argued over a generation ago, “The respondent is a human being. Whatever value the scholar may attach to the enlightenment he hopes to acquire through . . . an interview, he also ought to place value on the dignity, privacy, and courtesy of the person who has granted that interview” (Dexter 1970, 14). We need add no more on this point.

We also wish to highlight that what makes interviews so challenging can also make them enjoyable. In comparison to field experiments or surveys, the human interaction involved in interviewing is often less structured and more open-ended. Both the substantive and interpersonal possibilities are less bounded, and respondents have more power to shape the dynamics and outcomes of interviews. These characteristics can make interviews more unpredictable, demanding that researchers instantaneously devise adequate responses to unanticipated changes in the interview dynamic or setting. Yet interviews also allow researchers and respondents to develop greater rapport – such that a single exchange lasts for hours and reveals tremendous new insights into the research question. Given the great intellectual and personal rewards of conducting interviews, we expect this technique to remain the essential mainstay of the field research repertoire.

## Site-intensive methods: ethnography and participant observation

Political science has yet to embrace ethnography and participant observation wholeheartedly.<sup>1</sup> The very terms may seem foreign to some in the discipline, resonating with other fields, particularly anthropology, in which they are more centrally employed.<sup>2</sup> Ethnography may even have a certain exoticism, perhaps seeming like an almost mystical practice that some very different kind of scholar might use studying the culture of one small village in a remote corner of the Global South for two or three dogged years.

And yet, many political scientists have employed these methods fruitfully for generations. Fully 34 percent of all field research projects reported in our faculty survey included these techniques (almost always in combination with other techniques). To put this figure in perspective, it is slightly larger than the proportion of projects that employed survey research, and more than two and a half times the proportion of projects that drew on focus groups. According to our survey data, ethnographic techniques are used by political scientists from all subfields, though they are especially prevalent among comparativists. They have also been used in all regions of the world, though somewhat more commonly in Latin America, Africa, the Middle East, and Asia than in the United States, Canada, and Europe.<sup>3</sup> These techniques do, however, display a notable gender gap – the only techniques covered in our faculty survey that do.<sup>4</sup> Our survey data also suggest that the use of

<sup>1</sup> Some elements of this chapter have been adapted from Read (2010).

<sup>2</sup> When explaining that she used ethnography in her fieldwork, one scholar described it as “closer to methods used in anthropology” and “not mainstream in political science.” Interview, LM-12, September 6, 2012.

<sup>3</sup> Even among field research projects within the United States, 29 percent employed ethnography or participant observation. The corresponding figure for projects including European countries is 27 percent. For Latin America and the Caribbean, Sub-Saharan Africa, the Middle East, and the Asia-Pacific regions, the figures are 46 percent, 53 percent, 47 percent, and 51 percent, respectively.

<sup>4</sup> Site-intensive methods were employed in 44 percent of projects reported by women, as compared to 29 percent of projects reported by men. Pearson Chi-squared tests on the overall sample and on subsamples of United States-only and international projects find this gender difference to be significant at the .02 level or less.

ethnography and participant observation has stayed remarkably constant since the 1960s. Indeed, some of the best-known applications of these methods in political science – such as James C. Scott’s work in a Malaysian village and Richard Fenno’s study of US House members’ behavior in their home districts (discussed in more depth below) – took place long ago. Indeed, as we have noted previously, the type of immersive engagement these techniques entail has been and remains implicit in some of the mainstays of research and funding in the discipline, for example, in the Congressional Fellowship program, Fulbright funding, and so on.

The fact that scholars almost always use ethnography and participant observation in combination with other data-collection techniques suggests that many scholars incorporate insights gained through the employment of these techniques in projects not specifically, or exclusively, designed around them. That is, scholars who use these techniques often do so as part of a process of triangulation, a principle we suggest underlies good field research. Moreover, in conducting interviews for this book, we found that some political scientists hesitated to self-identify as “an ethnographer,” even as they asserted the value of in-person observational research. Frequently these scholars expressed uncertainty about the definitions of ethnography and participant observation, and whether their research fitted into those categories. This lack of conceptual clarity and hesitation suggest to us that the figures quoted in the previous paragraph may underestimate the use of these techniques. Using such techniques to at least *some* degree – even if sporadically – seems almost inevitable when conducting field research. Indeed, doing so is a quite natural aspect of engaging with context, a practice that most field researchers value, and which we also consider a principle of effective field research.

Perhaps in growing recognition of these dynamics, political scientists situated in multiple subfields have of late initiated an important dialogue about how ethnographic research approaches fit within our discipline.<sup>5</sup> These scholarly discussions have been facilitated by the formal institutionalization of a Political Anthropology and Sociology section of the Midwest Political Science Association, as well as the attention given to ethnography in several

<sup>5</sup> A recent edited collection by Edward Schatz (2009c) provides an outstanding overview of multiple perspectives on ethnography. Wedeen (2010) deftly surveys both the tensions and the complementarities that arise from the use of ethnography in political science. Other recent discussions include Bayard de Volo and Schatz (2004) and Wood (2007). Older treatments include Fenno’s reflections on his method (1986, 1990). In addition, Timothy Pachirat has written a book (forthcoming at the time of this writing) titled *Ethnography and interpretation* (2015).

short courses held at the APSA annual meeting, and at the IQMR, which takes place at Syracuse University each summer.

We believe it is essential to continue and deepen these dialogues. Clearly, debates exist among political scientists about the meaning and value of ethnographic methods. Even ethnographers disagree among themselves. Many scholars highlight the epistemological differences between interpretive and positivist ethnographers (Kubik 2009; Schatz 2009b; Wedeen 2010). Where interpretive ethnographers challenge the possibility of objectivity and seek to maximize their subjective understandings of everyday practices, positivist ethnographers employ ethnography as a method of uncovering a knowable empirical reality, the invisible truth. Additional debates question whether ethnography can or should be used to establish causality or attempt empirical generalization (Schatz 2009a; Wedeen 2010). And political scientists disagree about whether ethnography can be used in combination with other methods, or even other analytic approaches such as formal modeling and statistical analysis (Laitin 2003), or whether a “stand-alone, intrinsic-value political ethnography” model – dedicated to the production of contextualized knowledge – yields more value to our understanding of politics (Schatz 2009a).

In this chapter, we hope to offer some ideas to advance these dialogues. We intentionally include and discuss multiple approaches to ethnography. Our survey and interview research suggests that political scientists do ethnographic research in a variety of ways, and no “pure” template exists. We contend that political science will benefit from a better understanding of how these techniques fit into our discipline, how insights gained from ethnography and participant observation contribute to the study of politics, and what challenges are entailed when engaging in these techniques.

In this chapter, to facilitate exposition as well as to highlight their commonalities, we group ethnography and participant observation together as subcategories of what we call site-intensive methods (SIM, defined below).<sup>6</sup> We begin by clarifying what these techniques mean in practice, and provide a tour of the distinct lineages from which they emerge. We demonstrate that they can take many forms, and contribute to many kinds of research programs, both alone and in conjunction with other methods. In the subsequent section we show that, contrary to the impression some may have, site-intensive methods are well suited for building and testing social science theory, as well as for producing thick, descriptive narrative accounts. The rest

<sup>6</sup> Ben Read (2010) introduced this term in earlier work, and it is further developed here.

of the chapter highlights the issues and challenges of preparing for and using site-intensive methods in the field and provides strategies for resolving some of the difficulties researchers encounter when employing these techniques. Throughout, we draw on a wealth of examples from scholars in political science whose research has benefitted from site-intensive work.

## **What do SIM mean in practice?**

This chapter addresses a range of methods that includes various forms of ethnography and participant observation. We believe that, rather than drawing a line between these two traditions, or (as is more common) discussing or working within one with little or no reference to the other, political scientists have much to gain by considering them as parts of a single rubric, which we term “site-intensive methods” (SIM). To begin with, both ethnography and participant observation crucially involve observation. The researcher is not only asking a respondent to talk about an issue, event, or set of behaviors, but is also observing directly the dynamics or behaviors of interest. Despite the conventional distinctions made, we believe that both ethnography and participant observation involve a significant degree of participation by the researcher in the field site.<sup>7</sup> Whether officially and obtrusively, or perhaps without any formal recognition, the researcher participates to varying degrees in the social, economic, and political life of the field context being studied.

Furthermore, despite lively debates in anthropology and sociology that have problematized and essentially overturned the traditional notion of ethnography as a very long stay in one small (often rural and developing-world) community (Gupta and Ferguson 1997a), this stereotype appears to persist in political science.<sup>8</sup> We suggest that both older and newer, narrower and broader conceptions of this general form of research can be thought of as part of a common category of site-intensive methods. The introduction of

<sup>7</sup> In seeing the two techniques as substantially overlapping, our position is similar to that of Wedeen, who notes that “the term ‘participant observation’ is often used as shorthand for the double nature of these [ethnographic] activities, in which a researcher is both an actor and a spectator” (2010, 257). Others take a narrower view of what counts as participation, reserving it for situations in which the researcher joins or takes on the role of the research subjects: for example, becoming a social worker to observe social workers’ relationships with their clients, or joining the staff of a political campaign in order to understand campaigns (Brown 1986, 171).

<sup>8</sup> One senior researcher contrasted “pure number-crunching” with “pure ethnography in a village under the baobab tree.” Interview, LM-22, October 2, 2012.

this term is not meant to ignite controversy – just the opposite. We hope adopting it will help political scientists to put aside the conceptual baggage associated with ethnography or participant observation, facilitating a conversation about the very important characteristics and practices that unite these methods.

We define site-intensive methods as the collection of evidence from human subjects within their ordinary settings, where the researcher's interaction with the surroundings (as well as his specific questioning) informs the study. We do not conceptualize SIM in black-and-white dichotomous terms, but more as a continuum of the intensity of immersion. Still, at minimum, using SIM implies cultivating a deeper engagement with a locality, context, or set of informants than might occur, for instance, in a project based purely on standard one- to two-hour interviews. At the most basic level, scholars who engage in SIM observe people's everyday practices – they watch them doing what they would be doing even in the absence of the researcher (although, as we emphasize, the researcher's presence must generally be assumed to affect the site and the subjects in some way). SIM encompass forms of field research that are less structured by the researcher than are other techniques, entailing a high degree of contingency or open-endedness. At least to some extent, one waits for information to present itself rather than insisting on immediate answers. As one researcher commented in discussing her site-intensive research techniques: "The key to field research is letting it come to you as opposed to grasping really strongly and tightly at what you think you need to find."<sup>9</sup>

Most field research involves at least a certain degree of immersion in – or engagement with – a locality, broadly construed. Even if a scholar spends her days in the vaults of an archive, she is still likely to be browsing the daily press, catching portions of radio programs, and conversing with local colleagues. If "interviewing" rather than "observation" is a researcher's primary data-collection technique, he still cannot help but take cues from the surroundings: the home, office, or other locale where an interview is held; phone calls a subject takes that interrupt his conversation with the researcher; interactions with other people who happen upon the scene; or documents the subject provides. All of these cues open a window onto the worlds of subjects: they are *data* that help a scholar interpret the information garnered through interactions.

<sup>9</sup> Interview, LM-20, September 20, 2012.

Moreover, interviews might lead to follow-up visits, invitations to participate in relevant meetings or events, and so forth, leading to further immersion. For example, at the conclusion of an in-depth interview with a federal official about tribal-state relations, MacLean was invited to stay and participate in a conference call between the federal agency and state government officials on how to implement a new set of bureaucratic rules with tribal authorities. Another researcher recounted how, after the ice was broken during formal interviews with law-enforcement officers, it became possible to observe their work in real time and listen to conversations among them: “I didn’t have to ask [permission] after a while. I drove around with them on patrol, doing whatever they were doing. This changed the interviews from semi-structured to a free-flowing format.”<sup>10</sup> As interactions multiply and deepen, the attentive researcher naturally slips into a mode of flexible receptivity that goes beyond a simple question-and-answer format. Thus, there need not be a solid line demarcating a boundary where other methods end, and SIM begin.

While our intentionally inclusive conceptualization of SIM might seem to suggest that everyone is using these methods, of course not every political scientist does so. Yet many scholars who conduct field research do. We hope our bringing SIM’s prevalence – the use of site-intensive techniques both by those who employ them deliberately and by those who use them casually – into the spotlight will contribute to encouraging a productive dialogue among field researchers of all types on how to engage most effectively in such practices.

Perhaps inevitably, due to the broad use of these methods, site-intensive research can take diverse forms in political science. Projects involving SIM may vary in terms of the length of time spent in any one site, the number of sites studied, and the types of context in which SIM are carried out. Some site-intensive work strives for a high degree of depth in a single locale; other studies pursue shallower immersion in multiple sites.<sup>11</sup> Naturally, this kind of study can proceed in many different types of settings as well, from remote and rural communities to cosmopolitan and urban neighborhoods, from impoverished agrarian villages to astronomically wealthy, post-industrialist enclaves.<sup>12</sup> William Reno (1998) observed warlord politics in rural and urban

<sup>10</sup> Interview, BR-10, August 21, 2012.

<sup>11</sup> See Gottlieb (2006) for a useful examination of the tradeoffs between depth and breadth.

<sup>12</sup> Bayard de Volo (2009) candidly discusses her contrasting experiences conducting one ethnographic project in US casinos and another in post-war Nicaragua.

Sierra Leone while Joe Soss (2000) investigated poor peoples' everyday experience of the American welfare system. SIM can even include research where the "site" in question is not a physical location but a more abstract space such as an online community or a transnational network.<sup>13</sup> According to Hardin and Clarke (2012, 3), ethnographic field sites are not restricted to small-scale and localized spaces but include "expansive domains that connect localized and international spheres." Just one example would be political anthropologist Ayça Çubukçu's research on the World Tribunal on Iraq, which consisted of a series of meetings held in several locations over the course of two years (2011).

Despite the fact that SIM may take on many different guises – and even though opportunities for observation or insight often arise seemingly by chance when a scholar is conducting field research – we wish to emphasize that site-intensive methods are not haphazard. Preparing in advance to deploy these methods, consciously considering the many different examples provided by other researchers, and reflecting on one's position in the field site all help improve the use of site-intensive methods. Moreover, SIM can be – and often are – carefully employed in a supplementary role alongside other methods. That is, a researcher may build a project entirely around these techniques, but doing so is a choice, not an imperative. In short, there is no single template for this kind of field research, and political scientists can incorporate SIM into their own projects in a variety of ways.

Throughout this chapter, we discuss in depth two canonical books in political science – James C. Scott's *Weapons of the weak* (1985) and Richard Fenn's *Home style* (1978). We focus on these studies for three reasons. First, these milestone works testify to the multiple forms that ethnography and participant observation can take. The comparison of Scott and Fenn illustrates the striking variety in the length of immersion, number of sites, and types of context involved in site-intensive methods. Second, both books are probably familiar to many readers as archetypes of site-intensive methods. Finally, both authors explicitly detailed the rationale for using these methods and the techniques they employed in order to do so.

First, James C. Scott's *Weapons of the weak* draws on a project that consciously hews close to the classic anthropological model of ethnographic

<sup>13</sup> See Gupta and Ferguson (1997a) on the need to revise the "tradition" of ethnography in anthropology in order to study many questions that are global or transnational, where the "site" crosses multiple geographic boundaries. See Staple (2012, vii) on how "evolving global connectivity" blurs the boundaries of field sites and creates new ethnographic social spaces.

field research. Scott describes his book as a “close-to-the-ground, fine-grained account of class relations” in a Malaysian village, population 360, to which he gave the pseudonym “Sedaka” (Scott 1985, 41). He states that he spent at least 14 months in this hamlet, interviewing, observing, and taking part in village life (1985, xviii, 46). Scott provides a forceful explanation for choosing this mode of research: “It goes without saying that I have thought it important to listen carefully to the human agents I was studying, to their experience, to their categories, to their values, to their understanding of the situation” (1985, 42). While framed as an argument specifically for the importance of investigating class consciousness up close rather than deducing it from Marxian theory, his reasoning builds to a more general defense of research that attempts to understand the conflictual and contradictory world of human subjects first-hand and in those subjects’ own terms.<sup>14</sup>

The other well-known political study involving in-person immersion looks not at peasants in Southeast Asia, but at elite politicians elected to national-level legislative office in the United States. Richard Fenn’s *Home style* was motivated by questions concerning the relationship between politicians and those they claim to represent: “What does an elected representative see when he or she sees a constituency? And, as a natural follow-up: What consequences do these perceptions have for his or her behavior?” (1978, xiii). Fenn’s approach to answering these questions was to spend time in the company of members of the US House of Representatives in their home districts. He famously characterizes his research method as “largely one of soaking and poking – or just hanging around,” and situates it explicitly within the tradition of participant observation as practiced by sociologists and other political scientists (1978, xiv, 249, 295). (Ethnography is not mentioned.) In the text of the book and its long methodological appendix, Fenn candidly and rather self-deprecatingly explains his modus operandi of accompanying politicians wherever they would let him tag along, building rapport, recording their remarks, and asking questions when possible.

Clearly this was a far “thinner” form of engagement with a research milieu than was Scott’s village study. Relative to a single-site project, Fenn traded depth for breadth, studying eighteen different representatives and thus obtaining substantial variation on characteristics such as party affiliation and seniority (1978, 253–254). The total time he spent with each representative ranged from three working days to eleven, averaging six (1978, xiv, 256),

<sup>14</sup> On the kind of field research that *Weapons* represents, see also the discussion in Wood (2007, 127–129).

and on some of those days the research subject was available only part of the time. From this comparison, we conclude that site-intensive research can take different forms. That diversity, we posit, results from thoughtful decisions by researchers about how best to answer their particular research questions given what they already know about the context they are studying. The next two sections illuminate how Scott, Fenn, and others have done much more than “just” soaking and poking in the field.

## How ethnography and participant observation contribute to theory

Employing SIM can help scholars to accomplish many important analytic tasks. In fact, we would contend that obtaining valid data on some topics in some places may *require* a form of site-intensive research. The following six points highlight the sorts of research tasks, and the kinds of research topics, that SIM are particularly well positioned to address.

### Generating new questions and novel hypotheses

One fundamental purpose for ethnography in political science is the inductive generation of new questions and novel hypotheses. For Fenn (1978), the choice to conduct field research by “soaking and poking” was in part driven by his conceptualizing the project as blazing a trail through wholly uncharted territory. He argues that political scientists had previously all but ignored representatives’ understandings of their districts and how they behaved there (1978, xiii). Given this theoretical *tabula rasa*, Fenn contends that a “totally open-ended and exploratory” approach was needed:

I tried to observe and inquire into anything and everything these members did. I worried about whatever they worried about. Rather than assume that I already knew what was interesting, I remained prepared to find interesting questions emerging in the course of the experience. (Fenn 1978, xiv)

Throughout the book, Fenn highlights his openness to new questions *and* to new theoretical explanations for the outcomes he observed. Ultimately, this open-ended approach did yield new theories about why politicians and voters act the way they do, even if Fenn soft-pedals them somewhat. The new hypotheses he advances include: House members in the early phases of their careers are most attentive to their districts; members focus most on constituents who are well organized and thus accessible in groups; and

constituents care about personal attention, respect, and the assurance of access at least as much as they care about their representative's congruence with their own issue positions (Fenno 1978, 215, 235, 240–242). According to the author, none of these theoretical factors had been highlighted by existing scholarship that had employed other research methods.

Fenno's insights are unquestionably theoretically rich. However, the way in which he discusses his data-collection techniques could unwittingly and unnecessarily diminish the perceived value of site-intensive methods in political science. Fenno's use of the term "soaking and poking" and his heavy emphasis on the unstructured nature of his project suggest that his approach was necessarily speculative and free-form. To be clear, soaking and poking is indeed a productive and sometimes essential way to familiarize oneself with a research site and develop questions; many of our interviewees, including some oriented toward positivist modes of research, talked about its value.<sup>15</sup> But it is equally important to point out that site-intensive methods can and often do depart from this type of research: the unobtrusive observation of the dynamics of a field site is often far more theoretically driven and systematically reflective than what Fenno depicts. Moreover, while site-intensive methods do have enormous value in the beginning phases of a study, they can be usefully employed throughout a field research project. SIM can also be highly useful where theoretical classifications and hypotheses are already well developed and entrenched. In fact, Fenno's own acknowledgments of previous research suggest that he was engaging with a more well-established field than he admits.

### **Defining and selecting cases**

Site-intensive methods can also contribute to theory by allowing scholars to obtain critical information, often lacking from existing datasets, that helps to inform case selection. To give just one example, ethnographic observations may assist scholars in understanding the parameters of the broader population of potential cases. MacLean drew on ethnographic observations to expand the range of individuals who would be counted in the total population of village residents for a survey. On informal walks with neighbors to visit their farms, she realized that a small, but significant, number of people lived in isolated houses close to the fields, but far from the central village.

<sup>15</sup> For example, interviews BR-1, June 1, 2012; BR-2, July 30, 2012; BR-5, August 13, 2012; BR-6, August 14, 2012; BR-16, October 31, 2013.

The research team had missed these peripheral housing units in the initial map and census of the village, but local people confirmed that these mostly poor, non-indigenous migrants sent their children to the village school, shopped at the village market, and contributed to village funerals. As such, it was appropriate to consider them as belonging to the village community. Without the use of site-intensive methods, the survey sampling would have been systematically biased.

Another scholar's experiences reveal how site-intensive methods can be combined with other types of data collected during multiple trips to inform case selection.<sup>16</sup> After a pre-dissertation trip, a graduate student, who initially hoped to conduct field research in several districts in two states in India, chose instead to carry out his study in only two neighboring districts in the one state where he had more extensive language ability. During his second visit, a long-haul trip on which he began to collect his dissertation data, he used published government sources to create a census of village communities that had implemented the policy of interest and then randomly selected a subset of villages from that list. The graduate student then visited these randomly selected communities as well as many other neighboring communities, drawing on his interviews, informal conversations, and observations to match each in-sample village with one that was similar in all other aspects except that they had not yet implemented the policy investigated in the student's dissertation. Each of the above examples demonstrates the value of data gained from site-intensive methods for defining and selecting cases for further study, reinforcing an earlier point that SIM are often used in combination with other data-collection techniques.

### **Developing concepts**

Site-intensive research can also help scholars develop conceptualizations of key phenomena. Due to the value placed on parsimony in our discipline, many researchers feel pressure to reduce empirical complexity into stark abstract concepts – as when great heaps of state machinery are categorized dichotomously as either democracies or autocracies.<sup>17</sup> Such constructs may

<sup>16</sup> Interview, LM-11, August 31, 2012.

<sup>17</sup> See valuable work by Giovanni Sartori (1970), Robert Adcock and David Collier (2001), Gary Goertz (2006), and David Collier and John Gerring (2009) on the process of concept formation.

seek to capture macro-level phenomena such as regime types, relationships between states, and social classes, as well as micro-level phenomena such as party identification, acts of participation, identities, and interpersonal relationships. Other scholars develop complex conceptual apparatuses in an effort to retain more of the nuances of the underlying world.<sup>18</sup>

Either way, disconnects, incompatibilities, and slippage are bound to occur between concept and reality as the empirical world on which these abstractions are based is intensely elaborate. Cracking open the history of some of the most critical terms we employ reveals a labyrinth of competing definitions, contested measures, and shades of gray. One researcher may find that observed reality falls neatly into the categories that she devised or borrowed from other scholars; another may find such a matching frustratingly elusive and be more eager to analyze how the classification scheme itself was constituted and contested over time (Wedgeen 2008, 80–81). Sometimes the gap between accepted concepts and observed reality, always present to a greater or lesser degree, is minimally relevant – what statisticians might refer to as “random noise.” Yet on other occasions these gaps can prove critically significant and can spark innovation – a rethinking of what were previously settled conceptual categories.

Site-intensive research provides special opportunities for such innovation. It allows and even forces the investigator to confront the tensions between social science abstractions and the lived world of the people under study in a particularly vivid and immediate way. For example, Frederic Schaffer (1998) spent 14 months in Senegal studying what the concept “democracy” means in that context by examining how people use the word. Repeated visits to the Chinese countryside (along with other sources) gave Kevin O’Brien and Lianjiang Li the material from which they developed their concept of “rightful resistance,” referring to forms of collective action that draw on the very norms that power-holders articulate (O’Brien 1996; O’Brien and Li 2006). Prolonged study of a cluster of thirty villages in one region of Senegal gave Dennis Galvan insight into the “institutional syncretism” that resulted from the collision of precolonial society and a bureaucratic state (2004). In sum, this kind of intensive research is particularly valuable for refining, correcting, and even creating the core concepts we use to make sense of the political world.

<sup>18</sup> Examples might include Mark E. Warren’s associational typologies in *Democracy and Association* (2001), and James C. Scott’s classic discussion of patron-client ties (1969).

## Obtaining hidden data

The data that we need to study many political topics are inaccessible or *hidden* to a greater or lesser extent.<sup>19</sup> Information relating to some topics is intentionally hidden – sensitive, confidential, or even taboo. Powerful people, organizations, or governments may prefer that it not be made known, while less powerful actors may fear the consequences of sharing such information. Unvarnished facts, narratives, and opinions may be purposefully kept behind barriers.<sup>20</sup> For instance, Cyrus Ernesto Zirakzadeh (2009) acknowledged the difficulties he initially faced in finding and talking with Basque nationalist activists in Spain. In another example, it was only by combining statistical analysis with interviewing and participant observation techniques that Soss, Fording, and Schram (2010) could discern how the operations and logics of the “neoliberal paternalist” American social policy governing the poor persistently produced racial inequality and exclusion. Pachirat worked undercover in a Midwest slaughterhouse in order to investigate the hidden politics of violent labor (2011). Indeed, the intentional veiling of information in which political scientists are interested often has political underpinnings: we can imagine, for example, that employers may not want to publicize their everyday practices for requiring citizenship documentation in the United States or Europe, where immigration is a hot-button issue.

Other data may be unintentionally hidden. It is famously difficult, for instance, to operationalize many key concepts such as “democracy” or “representation.” Governments in some countries simply do not have the capacity to gather the information necessary to (reliably) produce some of the key indicators that political scientists often use in cross-national research. To give some concrete examples, political scientists have emphasized recently the importance of studying informal institutions, rules that are not formally documented but shape political life in profound ways.<sup>21</sup> For example, Freidenberg and Levitsky (2006) draw on ethnographic research to reveal the importance of informal party organization in Argentina and Ecuador. They argue that an exclusive focus on the formal statutes of party organization leads to “a flawed understanding” of party politics in much of Latin America

<sup>19</sup> Jourde (2009, 20) argues that ethnographic methods enable political scientists to see “political relations and political sites that are generally unseen, or ‘unidentified,’ by mainstream political science but which are nonetheless meaningful for local political actors.” Referring to such things, he borrows Denis-Constant Martin’s term “UPO” or “Unidentified Political Object.”

<sup>20</sup> Political research on such things – “hidden transcripts,” preference falsification, acting “as if,” and so forth – is voluminous (Kuran 1989, 1995; Scott 1990; Wedeen 1999).

<sup>21</sup> See Galvan (2004), Helmke and Levitsky (2006), and MacLean (2010).

where extensive informal organizations – from grassroots soccer clubs to clientelist networks – recruit activists and deliver votes without official sanction or recognition by the party (2006, 179). In discussing her approach to analyzing the genocide in Rwanda, Lee Ann Fujii (2009, 12) notes that “the key is finding ways to detect these less visible, but no less important, sets of meaning.”

As these examples suggest, and as we discuss in other chapters when examining the challenge of accessing data using other data-collection techniques, the problem of data inaccessibility can plague scholars no matter whether they are researching in authoritarian and repressive environments, or liberal democratic ones. Likewise, it is not a problem that relates more closely to any particular research topic, or to studies of elites as opposed to ordinary citizens. The ubiquity of the challenge makes finding ways to address it critical.

One strategy is simply “cold-calling” – making inquiries about elusive information out of the blue. While this may work under certain circumstances, often it will not, particularly if the information is inaccessible due to its sensitive or controversial nature. Perhaps even worse, such cold-calling might result in a scholar being granted an interview but obtaining answers that are deceptive, simplified, merely the “party line,” or the safe response. She might blithely record such information, unaware that it differs from what the informant believes on a more private level. To go beyond such an impasse may require building trust, waiting to observe unguarded moments, or finding other ways of opening access to a broader array of performances and practices by the actors involved in a scholar’s area of interest.<sup>22</sup>

In short, one powerful argument for using ethnographic techniques is that in some settings, it may be impossible to obtain valid evidence without using them. Sometimes, building ties of trust and familiarity with one’s research subjects and the institutions in which they are embedded is necessary for research to proceed. For more positivist scholars, ethnographic techniques can help “peel the onion” to remove layers of biases in order to arrive at a more reliable approximation of the truth (Allina-Pisano 2009, 54). For more interpretivist scholars, ethnography helps to reveal the contestation over what constitutes the “truth” for different people in different contexts, facilitating the study of how meaning-making operates in everyday practice.<sup>23</sup>

<sup>22</sup> Lisa Wedeen (2008, 88–89) discusses the genealogy of the interpretive emphasis on performative practices in her analysis of identity formation in Yemen.

<sup>23</sup> See Kubik (2009) for an excellent comparison of positivist, interpretive, and post-modern approaches to ethnography.

## Testing hypotheses

A great strength of site-intensive research techniques is that they offer rich opportunities to collect information on the observable implications<sup>24</sup> of hypotheses that are under scrutiny in a given project. That is, they provide ways to gather data relevant to multiple hypotheses more or less simultaneously from many elements of the research context, and thus to evaluate the extent to which these hypotheses accord with empirical realities. Thus, in a given site on a particular afternoon, one might observe a conversation involving several people, who also periodically respond to knocks on the door, telephone calls, news reports on a television, and so forth. Many fragments of the conversation might provide leverage on hypotheses concerning the various individuals' attitudes and beliefs, or the power relationships among them, or their interactions with others. (Alternatively, an afternoon's observation might be judged to provide little or no valid evidence, if, for instance, a supervisor or other authority figure happened to be present, changing the way subjects would otherwise behave; SIM require constant judgments of this type.) By collecting many forms of information – assertions that are made, grudges that are nursed, imperious or polite expressions used, body language, comments by individual A that later are discussed by individuals B and C after A has left the room, etc. – from many sources, scholars using site-intensive methods thus have particularly rich opportunities to engage in triangulation.

The fine-grained accounts of villagers' narratives, rituals, insults, and struggles in Scott's *Weapons of the weak* can be read as a series of tests of the "false consciousness" hypothesis, i.e., "the assumption that the peasantry (proletariat) actually accepts most of the elite vision of the social order" (1985, 40). The village ne'er-do-well, Razak, provides one such test: "As a beneficiary of local patronage and charity, however reluctantly given, one might expect Razak to entertain a favorable opinion of his 'social betters' in the village. He did not . . . 'They call us to catch their (runaway) water buffalo or to help move their houses, but they don't call us for their feasts . . . the rich are arrogant'" (1985, 12). Other pieces of evidence appear in the poor peasants' perceptions of local misers (1985, 13–22); the celebration that results when a job-displacing combine bogs down in the mud (1985, 163);

<sup>24</sup> *Designing social inquiry* called for qualitative researchers to improve hypothesis testing by identifying and evaluating hypotheses' observable implications, of which even a single case may have many (King et al. 1994, 28–31). Whether researchers subscribe to or reject other aspects of this book's advice, this point helps to highlight the power of SIM.

and the conflict over the village gate that had protected villagers' paddy-hauling jobs (1985, 212–220).

In *Peripheral visions* (2008), Lisa Wedeen draws on 18 months of field research to reveal the "observable effects" and underlying logics of "performative practices" with regard to nationalism, a particularly complex topic (2008, 15). For example, Wedeen describes the way Yemenis gather routinely in the afternoons to chew "qat" and share intense discussions about politics with friends and even strangers. Drawing on these observations of everyday political participation, she challenges Benedict Anderson's theory of national identity construction by showing how national attachments are constructed in the absence of a strong state in Yemen.

Often, site-intensive research encourages the observation of physical space as well as speech and practices (Yanow 2006). For example, MacLean (2004) observed the implications of different patterns of intra-family reciprocity in the varied architecture of family housing in cross-border regions of Ghana and Côte d'Ivoire. Where the survey data collected on family reciprocity revealed cross-sectional patterns of social support at one particular point in time, the concrete history of building visibly demonstrated the changes over time in how families were organized to support each other. The continued use of older extended family houses in the center of the Ghanaian villages compared to the prevalence of long-established nuclear family villas in the Ivoirian villages supported the hypothesis that differences in reciprocity emerged from longer historical processes of state building, rather than recent changes in the political economy.

### **Understanding causal processes**

Site-intensive methods can also enrich scholars' understanding of causal processes. One way of thinking about how they do so is to suggest that they can help researchers to identify "causal process observations" – "insight[s] or piece[s] of data that provide information about context, process, or mechanism, and that contribute distinctive leverage in causal inference" (Brady and Collier 2004, 277). "Process tracing" is essentially the search for such observations. Some of the most prominent disciplinary discussions about using qualitative methods to identify causal processes have emerged in the context of case-based research using historical sources.<sup>25</sup> Yet ethnography and participant observation can

<sup>25</sup> For example, Chapter 10 of George and Bennett (2004) gives an overview of process tracing with an eye toward research based on case studies drawing on historical documents. See also Tarrow (2004).

help scholars to collect data that can be used for the same purpose (even if those using SIM may not think in these terms or use this vocabulary). Indeed, site-intensive methods may be *even better* suited for illuminating causal processes, as they allow for active probing by the researcher rather than relying on the analysis of extant sources.

Fujii (2009) uses site-intensive research – ethnographic observations and purposively selected life history interviews – to understand how peasants could commit mass violence against their own neighbors during the Rwandan genocide in 1994. Fujii resists imposing the assumptions of the “standard approaches” to ethnicity and draws on her understanding of local culture and knowledge to make alternative meanings visible. She refuses to sort actors into fixed categories of “perpetrator” and “victim.” Instead, her ethnographic research reveals how actors have varying performances of the “script” for violence. As a result, Fujii views the genocide as having followed a dynamic and contingent causal process.

By contrast, Tsai’s (2007) research on public goods provision in rural China provides an outstanding example of how ethnographic observation can be mixed with large-*n* quantitative methods to gain insight into causal processes. In this project, site-intensive observation facilitates the interpretation of a survey of 316 villages in four provinces. Regression models estimated on these data offer evidence that, *ceteris paribus*, villages that possessed certain kinds of social institutions – notably, temple associations and village-wide lineage groups – also tended to provide good roads and schools. In an effort to elucidate the causal link between these social institutions and the outcomes of interest, Tsai carried out a series of focused qualitative comparisons among at least nine villages, visiting each for between 2 and 20 days. She found that in Li Settlement, for example, leaders were able to rally their constituents on the basis of a common lineage and community spirit, facilitating collecting donations to pave a road; in Pan Settlement, by contrast, long-standing conflict among the sublineages impeded similar efforts.

Overall, site-intensive research methods include multiple approaches that make compelling contributions to theory development in the social sciences. In the next two sections, we explore the challenges of preparing for and carrying out site-intensive work in the field.

## Preparing to use site-intensive methods in the field

Several prominent political scientists engage in site-intensive research, and the opportunities for dialogue on conducting ethnography in political science

are growing. Nonetheless, site-intensive methods are perhaps the least well-known across the discipline. In this section we consider several issues political scientists face when preparing to use these techniques in the field: determining that SIM are worth the effort and time; finding ways to start the process of observational research; balancing depth of immersion against breadth of multiple sites; and managing to be in the right place (to observe) at the right time.

### Whether and how to begin

An important question to consider in any project is whether the time and effort involved in site-intensive work are justified, relative to what can be obtained through (for instance) interviews. For some purposes and in some settings, a researcher can ascertain essential pieces of information from those with whom she needs to speak in single sessions. When this is the case, returning to an informant for a second or a tenth encounter may seem like a puzzling waste of time. Moreover, minimizing the time spent with each subject allows one to maximize the number of interviewees. Doing so facilitates triangulation, and may allow a scholar to obtain high degrees of variation on particular variables. If the scholar carries out a sufficiently large number of interviews, she may even be able to reap other benefits of large-*n* research.

But if SIM are indeed appropriate in a given project, one of the next questions concerns how to begin. How might one initiate a mode of inquiry that goes beyond (for instance) the cut-and-dried format and short duration of an ordinary interview, and instead involves observation that might extend over a protracted period of time? As previously noted, site-intensive research can evolve spontaneously out of other data-collection techniques. The investigator need not have at the outset – much less announce – a clear intention of repeating observational visits over a period of weeks or months. Rather, if people in the research sites accept his presence, he can simply continue to show up and assess, on a rolling basis, what is gained through site-intensive work; this is how Read proceeded in studying para-statal neighborhood organizations in Beijing and Taipei (2012, 287–290). Another approach is to embed oneself in some capacity – taking on a task or function that serves the community under study. For example, a researcher studying community-based organizations in Brazil identified specific ways to assist each organization's mission, for instance leading "train the trainer" English teaching classes and self-esteem workshops for youth. Doing so allowed her to be physically present at each organization for extended periods each day,

have casual conversations, and observe phenomena to which other data-collection techniques would not have given her access.<sup>26</sup> Soss conducted participant observation “in a shelter for homeless families, disability support groups, and welfare agencies” in addition to pursuing in-depth interviews (2006, 324).<sup>27</sup>

### **Immersion: depth vs. breadth**

Given that immersion is needed, how deep should that immersion be? We resist the notion that only very long-term immersion can be considered adequate: there is no minimum number of days or months that must be spent in the field in order for one’s work to “qualify” as SIM – and no amount of time in the field that *automatically* qualifies one’s work as of that type. Instead, we suggest that it is the quality of a scholar’s engagement with context while in the field that determines whether he has employed site-intensive methods. This view may conflict with the conceptualization of ethnography as a type of method that requires, indeed is almost synonymous with, lengthy field stays. And we of course agree that, as a general rule, scholars who spend more time actively engaging with a particular context are likely to come to understand it better than scholars who actively engage with that context for less time. Details, nuances, and dynamics of change are more likely to be overlooked by the latter; connections may not be made as deeply, nor relationships solidified as strongly. In some settings and for some purposes, trust between the researcher and her interlocutors may take weeks or months to build. Nonetheless, we believe it is counterproductive to set a bar that effectively excludes researchers who are arguably engaging in site-intensive techniques – hence our more permissive and inclusive understanding of those methods.

No matter how much time they decide to spend in the field in connection with a given project, scholars employing SIM face an inherent tradeoff between depth and breadth: between working more intensively in fewer sites or just one (observing more thoroughly and developing deeper relationships with informants), and carrying out research in more sites but spending less time at each. Many factors will inform scholars’ decisions on this front.

<sup>26</sup> Interview, DK-13, August 8, 2012.

<sup>27</sup> While we have emphasized the potentially fluid boundaries between SIM and techniques like interviewing, Soss notes that participant observation and interviewing can also be quite distinct enterprises with different strengths and payoffs (2000, 325–328).

The appropriate length of a particular field stay often emerges organically as the dynamics of the stay develop and a scholar learns more. Researchers employing SIM are particularly well placed to build trust when doing so is necessary, to patiently wait and watch for revealing moments or dynamics, and to come to understand the potential and possibilities of a field visit. Logistical concerns may come into play – for instance, the possibility and cost of transportation among sites and the need for multiple sets of accommodations the more sites are included. Likewise, intellectual imperatives will certainly need to be taken into account – for instance, one's language skills, and the similarities, differences, and relationships among the possible field sites.

One specific factor that might play a particularly important role in a scholar's decision about how many sites to study (and thus how long to stay in each) is the number of site-level independent variables she is interested in investigating, and how they vary across contexts. For instance, Scott chose Sedaka due to its apparent typicality, and because previous studies had established a baseline from which to assess change (1985, 90). While he does not explicitly justify his decision to focus on one village, we can surmise that long-term immersion in a single place was required to obtain access to "hidden transcripts," and that variation within the village itself (between elites and poor, and among different informants) was a more important focus than, say, interregional variation between Kedah and Johore. The overall framing of the research also seems to discourage the exploring of different sites: the denizens of Sedaka are intended to speak for peasants everywhere. Fenno adopts a casual tone in *Home style* when discussing how he selected his eighteen congressional representatives for study. He writes that he made "no pretense at having a group that can be called representative, much less a sample" (1978, xiv). But it *is* a sample, of course, and he makes clear that he selected it in such a way as to observe members and districts of different parties, regions, races, ages, levels of seniority, and electoral competitiveness. Thus, he pursued the familiar small-*n* strategy of obtaining variation on a number of potentially important independent variables. The problem of small sample size remained, but Fenno's logic is clear.

There are many reasons one might want to stay in a field site over a longer period of time. We have already mentioned that, in some contexts, trust and familiarity are required to obtain information.<sup>28</sup> As one of our interview

<sup>28</sup> The idea that repeat visits to a field site increase respondent trust and rapport was noted by many scholars we interviewed, e.g., interviews, LM-12, September 6, 2012; LM-13, September 7, 2012; LM-16, September 11, 2012.

respondents explained, these benefits do not require a scholar's uninterrupted presence in the field site. Instead, sometimes rapport can be increased by a scholar leaving a field site and returning a few weeks later, thus giving research subjects "time off from you."<sup>29</sup> Further, even the most cooperative and forthcoming of informants may "change their stories" over the course of a single interview and between interviews. Scholars interested in tracing that variation will do well to interact with their subjects on more than one occasion. Just as different question wordings, question orderings, and contexts can lead to quite dissimilar responses in survey research, posing related questions somewhat differently in a series of interviews can result in respondents revealing different and possibly conflicting beliefs and experiences. Moreover, interviewing respondents in the presence of different groups of neighbors or colleagues (if one-on-one conversations are impossible) can lead to interesting response variation if their relationships condition what is revealed. Finally, from the perspective of a more positivist scholar interested in obtaining the "truest" or most objective measure of an individual's perspective, multiple interactions may yield data of higher validity. For more interpretive scholars who emphasize subjectivity, the benefit of multiple sessions may be the inclusion of a greater number of voices and perspectives.

Another factor conditioning the amount of time a researcher will spend at a site, or how many times she will return, concerns the degree to which data collection hinges upon the observation of particular events or dynamics. To the extent that one is essentially waiting around for special circumstances to crop up, the time investment will grow, though the payoff may make it worthwhile. Fenno clearly exercised considerable patience, traveling to each of his eighteen subjects' districts at times when the representative would be present there, then lingering in the background until moments when he became available. He seized upon moments of special insight, when the politician might let slip an unguarded remark or even confide in Fenno something he would ordinarily not reveal to an outsider.

Straus's *The order of genocide* serves as an example of artfully balancing many of the factors we have just discussed. He interviewed approximately 230 Rwandan genocide perpetrators (most of them randomly sampled). He also made trips to five specific locations scattered around the country, speaking to "a cross section of Rwandans, including survivors, perpetrators, current and former officials, and other local leaders" (Straus 2006, 5–6). The information he obtained through the site visits illuminated the specific

<sup>29</sup> Interview, LM-20, September 20, 2012.

processes through which catastrophic violence broke out in these communities, processes that turned out to differ from one locale to another.

In sum, multiple factors will condition scholars' decisions about how many field sites to incorporate into their projects, and what to do there. No single model will work for every project. Nonetheless, we hope to have provided some ideas about the considerations that it will be helpful to take into account as scholars chart their course in the field.

### **Being at the right place to observe at the right time**

How can researchers be sure to be in the right place to observe at the right time, particularly if their time and resources are relatively limited? Most researchers' time is constrained by competing professional or personal commitments. Furthermore, funding for field research is not always sufficient to support long-term stays. We raise here a few issues and strategies to consider in order to maximize the time available for observation in the field.

A shorter, preliminary field stay can help scholars to refine the research strategy they will subsequently employ. During such a field stay, a researcher can begin to introduce herself to potential research communities or participants, finalizing the details of her ethnographic visit once she returns to her home institution. Of course, the process of obtaining the access needed to carry out site-intensive work effectively will vary by context. In some places, researchers might need to gain official approval to engage in such work – for instance, from government agencies. In other contexts, such as public meeting places and events, scholars may be able to show up whenever they wish.

One issue scholars may need to consider as they plan site-intensive work is seasonality. Many political topics have a seasonal quality to them. For instance, for a question on political participation, the timing of the electoral cycle might be critical; for an analysis of philanthropy, the proximity to tax collection could be salient; and for an examination of poverty alleviation in Africa, the timing of what are commonly known as the hungry or lean seasons, or the occurrence of major religious holidays, might be important. Holidays and vacations could mean a complete absence of activity in a site such as an office or a legislature, but could provide special opportunities for observation in other settings, such as people's homes. Indeed, scholars should consider how they can "make use" of seasonality, particularly if their project includes more than one site. For instance, they may be able to observe one site when it is "in season," then switch to another site to make it their

“off-season” case. Of course, the seasonality of political topics also affects how researchers employ other data-collection techniques. But to take full advantage of direct observation in site-intensive work, researchers must think carefully about how different seasons might affect the presence, salience, or dynamics of the activity under study.

Another factor to consider is how the scholar wishes to position his visit vis-à-vis an “official” event of interest. Many scholars who use ethnography and participant observation have noted that the most revealing interactions can happen immediately after such an event ends. This is similar to the observation made by scholars who conduct in-depth interviews that respondents often gush with candid revelations and acute insights only *after* the recording device is turned off, or the notebook is closed. Researchers can often find good excuses to tarry and observe informal conversations as a group disperses. As always, every effort should be made to record notes on these observations as quickly as possible before one’s memory of them fades.

In the preface to *Ambiguities of domination*, a book written on the basis of two-and-a-half years of fieldwork in Syria, Lisa Wedeen writes that, although her research was built around formal interviews, and benefitted from them,

it was really the events of everyday life, the periods of “hanging” out – of drinking coffee, studying at the university, teaching exercise classes, making olives, babysitting, hosting salon-like gatherings in my home, going to films, shopping for groceries, riding the bus with friends to visit relatives in faraway places – that produced the chance encounters and enduring connections that animate this book. (Wedeen 1999, vii)

The leveraging of serendipity and the observation of the everyday, over time, into cumulative insights about the nature of authoritarianism under Hafiz al-Asad epitomizes the long-term payoffs that site-intensive study may yield.

## Issues and challenges of using site-intensive methods in the field

In addition to dilemmas about where and for how long to use site-intensive methods, political scientists face additional issues and challenges when employing these techniques in the field. Here we address just a handful: the ethics of observation, the nature of positionality or objectivity, compensation, and interpreting one’s experience in the field. While each of these issues may arise when deploying other data-collection techniques, they frequently pose more intense challenges when using SIM.

### **Ethics of observation**

We begin with the ethics of observation and issues around informed consent. Frequently researchers are more attuned to these issues when they are carrying out a survey, in-depth interviews, or an experiment. When using these other data-collection techniques, the researcher and participant can often point to a specific moment when they discussed the study, shared information about the nature of the encounter, and then agreed to proceed. But these questions come into play with SIM as well: what about a participant's ability to refuse to be observed?<sup>30</sup> This is a particularly delicate issue when the scholar has become so immersed that subjects become friends or forget that they are under study.

One of the ways that a researcher can facilitate a more meaningful process of informed consent for ethnographic observation is by ensuring that the project has been introduced publicly to the broader community. Precisely because site-intensive methods require scholars to immerse themselves in observing and understanding the everyday practices of the field sites, the meaning of informed consent frequently goes beyond an individualized contractual exchange and potentially involves the generalized awareness of a wider group of participants. Indeed, in some developing countries, no individual-level interactions can proceed until a village or neighborhood-wide public meeting has been held, where the researcher openly introduces the research team and the study objectives. Even after such a public meeting was followed by the village gong-gong beater passing the news at 5 a.m. in Ghana, MacLean's research team still had to dispel many myths and misunderstandings as they went door-to-door greeting and introducing themselves. Of course, this kind of face-to-face meeting may be neither logically feasible nor necessary in larger scale communities or advanced industrialized contexts. Nonetheless, many contexts have periodic large, public meetings organized for other purposes, in which a researcher may ask to make a brief introduction. MacLean's work in the United States provides another example. She was invited to deliver formal presentations of her work on Native Americans at large inter-tribal meetings as well as more informal updates at smaller tribal council and committee meetings. These introductions ensured broad awareness of MacLean's role as a participant observer in public meetings, stimulated approval of the study from existing study

<sup>30</sup> See Yanow and Schwartz-Shea (2008) on the need to reform IRB policy. The authors emphasize that IRB policy is not designed around ethnography, and that, indeed, ethnography may invert the traditional power relationships where the participants refuse entry to the researcher.

participants, and generated new contacts for additional in-depth interviews. In still other contexts, the researcher may be observing as many different people move in and out of the field sites, without any ready opportunity to introduce herself or to inform the subjects of the study. In these cases, scholars should think carefully about the potential sensitivities of these observations and whether and how to use them in published work without unwittingly harming subjects who were unaware of being observed.

None of the above strategies is useful, however, when the researcher is an unobtrusive or undisclosed participant. Indeed, in this type of project, the scholar has often purposefully chosen to use site-intensive methods in order to understand what is hidden or politically sensitive. As such, the project's success hinges on his ability to observe without subjects being informed fully, or at all, that they are under study. For example, Timothy Pachirat concealed his identity when he applied for a job in an industrial slaughterhouse in Nebraska. Pachirat (2009, 148; 2011) describes how his brown skin, status as a young, male, Southeast Asian immigrant to the United States, and previous manual-labor experience enabled him to be recognized and hired as "typical entry-level slaughterhouse material." For 5½ months, Pachirat worked side-by-side with other employees, sharing stories, eating lunch in the cafeteria, even arranging rides and running errands with one co-worker who became a friend. He never told the workers or management that he was actually a graduate student at Yale University conducting doctoral research.<sup>31</sup> Pachirat recounts his initial paranoia and sense of danger that management might discover his identity.<sup>32</sup>

In these instances, the scholar needs to critically reflect on where he positions himself on the spectrum of ethical behavior presented in Chapter 4, what he feels constitutes an acceptable level of deception, and what the implications of the chosen level are. In the example above, Pachirat was thinking reflexively about the tradeoffs involved with different strategies of access – specifically, how direct access as an undisclosed entry-level worker, as opposed to formal access (with management approval) or proxy access (via interviews outside of the slaughterhouse) would shape what he saw and how he interpreted that information. Pachirat demonstrates the value of

<sup>31</sup> Pachirat is explicit that his intention was not to produce an exposé of a particular company, its management, or employees, so he does not report the names of the company or people he met in his published work.

<sup>32</sup> See also Pachirat's interview with James McWilliams, a writer on food politics and animal rights, at [www.james-mcwilliams.com/?p=1577](http://www.james-mcwilliams.com/?p=1577), accessed March 12, 2014.

continuing to think reflexively throughout the project, and not just at the initial design stage. He describes how, during the course of his participant observation, the benefits of direct access began to be outweighed by the ethical costs when he was promoted to a quality control position, and thus charged with monitoring and enforcing the slaughterhouse standards. At that point, he decided to quit:

Ultimately, the ethical dilemmas inherent in the work of quality control – both the diversion and redirection required in relating to federal meat inspectors and the surveillance and discipline I imposed on subordinate workers in the plant – became, in my judgment, untenable, and I resigned from my position and left the slaughterhouse in December 2004. (Pachirat 2011, 156)

During the rest of Pachirat's field research, he relied on proxy access, observing and conducting interviews as a known researcher with ranchers, slaughterhouse owners and workers, inspectors, and community and union organizers. While the knotty issue of tolerable deception comes up with other data-collection techniques, it may be particularly intense in site-intensive methods and, as we discuss in Chapter 9, in field experiments.

### **Positionality issues**

Researchers using site-intensive methods must also consider how their positionality vis-à-vis study participants influences the data collected. These issues again affect the employment of other data-collection techniques as well, as we note in our broader discussion of positionality in Chapter 4. Nonetheless, the dynamics may become more intense and readily apparent when the researcher is deeply immersed in a particular field context and in consistent contact with study participants. Indeed, both advocates' conception of the primary contribution of site-intensive methods and skeptics' dominant critique thereof center on how the intimate positionality of the researcher in the field site may either facilitate or, alternatively, undermine the validity of the data collected. The implicit (and erroneous) assumption in both of these perspectives is that somehow a scholar using site-intensive methods necessarily becomes an "insider" over time due to their intensive immersion in the field sites.

But using site-intensive methods does not erase positionality or the importance of the scholar's identity and how it is perceived. Drawing on the literature in anthropology and elsewhere (McCorkel and Myers 2003; Schatz 2009b, 6–9), we highlight that a researcher's positionality is constantly

shifting and being renegotiated in the multiplicity of contexts discovered in the field. Since any scholar's identity is socially constructed, multivalent, and dynamic, immersion in new cultural and geographic spaces can spur transformations in how the researcher speaks, moves, acts, and is perceived by others in the field sites over the course of a project (Hardin and Clarke 2012). While this is true for any researcher, those using site-intensive methods have a chance to reflect upon and explore the problems posed by their positionality in a deeper way than when using data-collection techniques involving one-shot interactions with subjects. These scholars gain more information through repeated interactions and engagement with context, which can help provide perspective on the dynamics of positionality and how it is shaping the research process.

Indeed, sometimes social differences, or being identified as an outsider, can smooth entrée into a community and enable a scholar to obtain "insider" meanings (Horowitz 1986; Tamale 1996).<sup>33</sup> Fujii (2009, 34–35) found that her status as an obvious outsider in Rwanda allowed her to pose questions that would have made interview respondents anxious or suspicious had they been asked by an insider. In one example, it was only after Fujii's translator reassured a woman that the questions she was being asked were coming from Fujii that the woman calmed down and patiently explained certain events' underlying logics – logics that would have been obvious to most native Rwandans. Power differentials are to some degree inverted in such interactions, as subjects' superior expertise on the research topic is acknowledged. Indeed, this acknowledgment can be extended past the specific domain of the research study to other situations of everyday life. Finally, while a certain amount of cultural competence and knowledge is critical to gain a base level of trust, often an outsider status facilitates open dialogue with multiple groups that may be in persistent conflict with each other. The researcher's outsider status can even be viewed explicitly by historically hostile groups as an opportunity to share information with each other indirectly and thus learn about the other group's views with a hope of improving the relationship.

It is of course impossible to identify the precise effect scholars' position and personal interactions have on the data they collect through SIM. Nonetheless, and regardless of whether scholars perceive their positionality

<sup>33</sup> A researcher who worked in a post-apartheid setting in Southern Africa noted that "because I was an outsider, they were more likely to trust me." Interview, LM-12, September 6, 2012.

to be a help or a hindrance, they should do what they can to critically reflect on their role in the field site and its impact on their data. One way scholars can do so is by merging their ethnographic field-note journal with any personal diary they may be keeping, thereby providing a more comprehensive record of the context for the observations. Scholars should also be attentive to opportunities to change their position and somehow level the playing field with the study participants. And even if differences remain, awareness or perhaps even acknowledgment of those differences by the researcher can build trust, potentially bridging or eventually even attenuating disparities.

### **Compensation and giving back**

Given the immersion that site-intensive research entails, researchers who use such techniques may build more extensive personal ties and networks of trust in the communities where they work, and with the study participants who occupy them, relative to those who use other methods. These relationships can facilitate a broader understanding of politics, but may also entail dense webs of obligation. As such, deciding how to “give back” to study participants and communities constitutes a significant challenge in site-intensive research. Indeed, scholars who engage in ethnography and participant observation may be uniquely aware of and affected by any inequalities or friction created between individuals, groups, or communities by the exchange of compensation or gifts.

Given the potential consequences involved, scholars using site-intensive methods often talk to other researchers who have worked previously in the same or similar field sites to find out what kind of compensation might be appropriate in the particular context. Anticipating these issues in advance might be particularly important if the most appropriate and valued gifts must be purchased elsewhere and brought to the field sites. An alternative strategy is to wait until later in the field visit, when the researcher herself has become more immersed and knowledgeable about local practices. Compared to other data-collection techniques such as interviewing or survey research, the quandary of compensation for site-intensive methods is not a one-time transaction with individual respondents, but is navigated over a longer time period simultaneously with individuals and a wider community of participants. This constant negotiation of what is expected and appropriate to give in the field site can be one of the most unexpected and draining aspects of field research. Some scholars describe the weight of the

accumulating obligations as “enormous guilt” felt during fieldwork, a feeling that does not diminish with one’s departure.<sup>34</sup>

Another way ethnographic researchers can avoid creating inequalities or friction while still giving back is to adopt a collaborative approach to research from the very beginning of a project. Involving participants in project development and implementation may make it more likely that the research will be relevant to, and promote long-term benefits for, the field sites. While we advocate for collaboration more generally in Chapter 4, such an approach may be particularly feasible for researchers using site-intensive methods due to their deeper involvement, often over longer periods of time, with those they involve in their work.

### **Interpreting and navigating ties to the field**

Ethnographic researchers frequently write up their observations of encounters, public events, and activities on a daily basis. As they do so, they often go beyond simply recording discrete facts or events: they interpret, compare, and process their reactions to what they observed. The explicit recognition that scholars using site-intensive methods are continuously reflecting and interpreting is what some scholars have termed an “ethnographic sensibility” (Schatz 2009b, 5) and others “an ethnographic or anthropologic approach to life” (Staple 2012, ix). According to Shehata (2006, 260), ethnographers’ reflexive scrutiny and analysis of their interactions with study participants generate more evidence about how those participants think and understand their social world. The scholars’ notes often include multiple voices and perspectives on any particular event, and their interpretations often draw on multiple sources. In other words, they triangulate and analyze as they write. Moreover, their deep understanding of the field context strengthens their ability to identify dominant narratives and weed out lies and self-serving patter. Critics of site-intensive methods may question their supposed lack of objectivity, yet scholars using this approach are systematically cross-checking, comparing data gained from different sources, and reading between the lines.

Just as scholars using site-intensive methods are intensely aware of how the field context shapes what they see in, and how they see, the world, they are often very self-conscious in their efforts to reimagine and reconstruct the

<sup>34</sup> Interview, LM-5, August 27, 2012.

field site when they continue to interpret and write after they leave the field.<sup>35</sup> Yanow (2009) encourages ethnographic researchers to employ the first-person singular “I” and include in their writing the details of the author’s presence – or evidence of “being there” in the field – for example, contextual information about the research setting, time, selection of participants, positionality, and the like. One of the challenges is how to communicate the value of the theoretical knowledge gained to social science audiences without sacrificing too much of the rich complexity and detailed understandings of the various field contexts and insider meanings.<sup>36</sup>

Relatedly, the deep integration experienced by scholars who engage in SIM also necessitates serious reflection about their long-term obligations to communicate and share their interpretation and writing after they have left the field. Communities under study may request follow-up engagement from the investigator because they wish to receive the benefit of the research findings, but they also may wish to remain connected for more personal reasons. One researcher expressed her regret about not maintaining such contacts: “I should have known to keep in touch with people better. I always knew I would return. I wish I could have found time to call every couple of months and say hello. When I returned, a good number of people were annoyed.”<sup>37</sup> One caveat is worth noting here, however. Some of the proposed strategies may be particularly challenging to implement for graduate students and young faculty who are at an early stage of their career. But the process of using site-intensive methods in the field can be navigated with ongoing dialogue between the researcher and subject communities about their mutual expectations and responsibilities.

---

## Conclusion

---

Ethnography and participant observation hardly lie at the core of graduate training in political science today. We have sought to show why these methods deserve much more attention than they have traditionally received

<sup>35</sup> See Yanow (2009) on the additional interpretive moments involved in “deskwork” and “textwork” by the ethnographic researcher and reader of an ethnographic text.

<sup>36</sup> Wedeen (2010, 256) highlights the importance of enabling a conversation among political scientists who use different research methods. Hardin and Clarke (2012) highlight the “paroxysm of conscience” that ethnographers experience as they attempt to reconcile the complexity of what they have learned in the field with the demands for knowledge from the academic scientific community.

<sup>37</sup> Interview, LM-5, August 27, 2012.

in our discipline. Much of the political world – motivations, perceptions, preferences – is kept hidden: organizations, politicians, and even ordinary people may wish to present a public face that differs from the inside story. Correspondingly, those who study politics can ill afford to turn their backs on approaches that entail deep engagement with context – one of the principles of good field research – and that are designed to lower the barriers between researchers and reality.

Further, these methods can be adapted to many purposes (hence their cross-disciplinary following). They lend themselves to interpretive modes of research that eschew some of the tenets of positivist social science. But, much more than is commonly appreciated, they can also help scholars with any epistemological approach make significant progress in a host of important analytic tasks, aiding with generating new questions, selecting cases, constructing concepts, establishing and testing hypotheses, understanding causal processes – and, thus, building theory. In addition, no matter what purpose these techniques are used for, employing them almost invariably involves significant critical reflection. Indeed, the field notes that scholars who use SIM typically take are *full* of the type of analytic reflections that we believe are a critical principle of good field research. In addition, given the existence of those very careful notes, scholars who engage in SIM are uniquely positioned to be extremely transparent about their research practices, helping readers to understand their research techniques clearly, and why and how they were employed. This research transparency is yet another principle that we believe underlies effective field research.

The purpose of site-intensive methods is to access information that cannot be obtained using other data-collection techniques. Scholars who use SIM believe that these techniques allow for the collection of highly valid data due to the deep relationships, rapport, and trust that often form between researchers and study participants when these methods are employed. Through prolonged interactions, the researcher aims to get below the surface and past the “party line,” and the vague or evasive or unconsidered answer. Watching people in the context of their everyday lives makes available data that could never be captured through a highly structured survey questionnaire or a single interview, useful though those data-collection techniques may be in other ways. Given their particularized and unique strength, scholars often use SIM in tandem with other data-collection techniques, meaning that they are often specifically employed as part of a strategy of triangulation.

Almost by definition, these methods involve an investment of time and effort that is greater than other methods on a per-informant basis. Yet

engaging in SIM need not be more “costly” *overall* in time or effort than employing other data-collection techniques discussed in this book. Full-blown immersion for many months at a single site is not the only form that these methods can take. Political scientists may have good reason to split their time among multiple sites, or to use SIM in limited ways to augment other methods.

Finally, we have aimed to articulate an encompassing conception of site-intensive methods as an ensemble of techniques with roots in various parts of the ethnographic and participant-observation traditions. We argue that interpretivists and positivists alike can make good use of these methods. To some political scientists, SIM may seem to lack objectivity and to be impossible to replicate. Though these techniques can be free-form and open-ended – and such research can be highly useful in exploratory studies – it is also possible to apply SIM in systematic ways. Bringing a carefully crafted set of hypotheses to site-intensive work can help the investigator to focus her attention on just those things that pertain. Moreover, as the discipline develops stronger norms around transparency in qualitative as well as quantitative research (as discussed in Chapter 11), ethnographers and participant-observers will be pushed to find ways to make their work more accessible to other scholars.

Survey research is a thoroughly conventional and seemingly well-honed tool in the political science work-shed. Numerous influential books and articles in the discipline have survey data at their core. It is enshrined in institutions like the University of Michigan's venerable Survey Research Center, together with its longstanding training programs. Doctoral programs commonly require coursework in the kinds of statistical methods used to analyze quantitative survey data, and in many departments it is common for Ph.D. students to base their work on such data.<sup>1</sup> To be sure, it appears that more dissertation writers and senior scholars conduct secondary analysis on existing aggregate datasets than field their own survey, and this has been the case for some time (Manheim and Rich 1986, 123). Still, original survey research conducted in the field has enjoyed a well-established place in the discipline since its introduction in the social sciences in the 1940s.<sup>2</sup>

A survey was used to collect data in 30 percent of the projects reported in the 2011–2012 FRPS survey – and this figure has stayed quite constant since the 1960s. Further, research within the United States was only slightly more likely to employ a survey than was international work.<sup>3</sup> Scholars in our discipline have conducted surveys in all regions of the world, with a somewhat lower incidence in Western Europe (22 percent of projects involving surveys include locations there) and a somewhat higher one in Sub-Saharan Africa (42 percent), perhaps due to variations in the availability of reliable survey data on political questions. And political scientists from almost all

<sup>1</sup> Schwartz-Shea found that 66 percent of a sample of fifty-seven top doctoral programs in political science had a quantitative/statistics methods requirement in the early 2000s (2003, 380).

<sup>2</sup> See Brady (2000) for an overview of the contributions of survey research to the discipline as well as a call for greater support for graduate students and faculty who have both “in-depth area studies knowledge” and training in sophisticated survey methods.

<sup>3</sup> Of projects with field locations only in the United States, 33 percent employed surveys, versus 28 percent of projects that extended to non-US locations.

subfields do survey research in ample numbers, including international relations, public administration and public policy scholars.

In many respects, political science is well equipped to teach researchers how to use this technique. Creating and implementing surveys for academic research requires a number of technical skills, which many standard textbooks, and courses based around them, aim to provide (Rea and Parker 2005; Leeuw, Hox, and Dillman 2008; Fowler 2009; Groves *et al.* 2009; Marsden and Wright 2010). Such texts provide general treatments of key elements of the survey process, from laying the overall design foundations to analyzing the data. In addition to theoretical background on topics such as sampling, they also have much advice to give about multiple kinds of substantive and practical matters that researchers confront in the field, such as minimizing interviewer bias and maintaining ethical standards with human subjects. In short, they constitute an essential starting point.

Vital though these published methodological works are, they convey only part of the wisdom that political scientists who seek to conduct a survey in the field need. Most do not focus specifically on political research, and they cannot do justice to the great variety of surveys that political scientists have devised and carried out. Moreover, they are generally written with applications in the United States (and to a lesser extent Western Europe) in mind. Thus they barely begin to consider the vast range of special circumstances and obstacles that field researchers encounter overseas, or even closer to home. Indeed, some of our interviewees explained ways in which standard prescriptions led them astray or had to be disregarded in the field. These textbooks also are not designed to address all the ways in which survey research might dovetail with other components of a research program.

Published work examining the methodologies of field research, for its part, often centers on modes of data gathering *other* than surveys. One recent overview of fieldwork in the authoritative Oxford Handbook series steered away from specific examples of surveys (Wood 2007). Though this essay placed surveys conducted in the field under the rubric of field research, defining the latter as “research based on personal interaction with research subjects in their own setting” may have shifted the focus away from surveys, which often involve only thin forms of interaction with subjects (Wood 2007, 123). While there are exceptions to this tendency to separate surveys conceptually from more deeply interactive forms of data collection (examples are discussed below, such as Sieber 1973), political scientists sometimes think of fieldwork and survey research as two different things.

In addition to these resources, political scientists who anticipate conducting a survey in the field have at their disposal, and can draw lessons from, the work of those who preceded them. Political scientists, sociologists, economists, anthropologists, and others have confronted the challenges of doing survey work in a wide variety of field settings for many decades (for example: Rudolph and Rudolph 1958; Bulmer and Warwick 1983; Dillman 2002; Converse 2009).<sup>4</sup> The record left by these many creative efforts is long, and implies plenty of practical expertise, sometimes conveyed in methodological appendices or chapter-length reflections on the research process.

However, often this expertise has not been compiled in ways that make it easy for today's researchers to access it. Indeed, even milestone works of political science based on surveys often contain only partial explanations of the fieldwork that underpinned the data collection. For example, *The American voter* included virtually no information regarding how the data upon which it was based were gathered; its preface simply stated that information such as the questionnaires and sample design "may be obtained upon request" from the Survey Research Center at the University of Michigan (Campbell *et al.* 1960, vi). A landmark of post-World War II comparative politics, *The civic culture*, did include the English-language questionnaire and an explanation of the sampling technique used in each of the five countries under study, but otherwise provided only a "very brief outline" of the process that led from the formulation of the original questions to the repatriation of the IBM punch cards containing the data (Almond and Verba 1963, 47). In this work as well, then, some of the central elements of carrying out surveys in the field, and the challenges involved, remain unseen in the background. Moreover, changes in societies and communications technology mean that even when details are shared, yesterday's lessons do not always apply in the present.

The purpose of this chapter is to fill the multiple gaps identified in the above discussion. We examine how the actual practices of survey research are woven into scholars' broader fieldwork endeavors, and stress how scholars' engaging with the survey context can encourage greater survey participation and enhance the quality of survey research more generally. We begin with

<sup>4</sup> For early perspectives from the heyday of the modernization school, see also the other articles collected in the Autumn 1958 issue of *The Public Opinion Quarterly*, titled "Special issue on attitude research in modernizing areas." Townsend, Sakunthasathien, and Jordan (2013) recount in detail the story of an ongoing survey project focusing on households and economic activity in Thailand; though rooted in economics, this contains many lessons for political scientists, particularly those working in rural, developing-world settings.

some reflections on the nature of political survey research and the many forms it can take in the field. We also highlight the many opportunities for constructive synergies between survey work and other kinds of data gathering in the field. We discuss multiple ways of using information collected through interviews or archival work, for instance, to inform and improve a survey, and to build upon and enhance findings from surveys through follow-up work. Indeed, as we observe, what could be seen as a boundary or tradeoff between the gathering of thin data on the one hand, and rich, intensive, thick data on the other, can sometimes be transcended. Rather than being mutually exclusive, these different forms of data can feed and inform each other in productive ways.

Next, we dig more deeply into the purposes for which and the conditions under which political scientists might contemplate fielding a survey. As part of this discussion, we address the various ways in which survey research contributes to furthering analytic goals in a project, and to building theory. We anticipate our treatment of this question here by noting that, in our view, surveys are remarkably versatile, with some caveats and limitations. Surveys are generally conducted in order to provide data for use in making descriptive inferences and testing hypotheses. A survey can provide a baseline understanding of the population under study – whatever that population might be – and this can be used to inform further stages of field research, for instance by selecting cases or generating hypotheses to be tested. A survey might even be used to clarify the concepts that are under study in a given project. Still, given the highly structured form that surveys take and their comparatively high cost in terms of time and money, it is more common for a survey to come after such basic elements of a research design as concepts and hypotheses have been largely or entirely settled. As we will again note in connection with experimental research, by the time a scholar takes on the complex process of setting up a survey (often in concert with collaborators), many key analytic choices and tasks will likely already have been made and carried out.

After probing the intersections between surveys and other techniques, and theory and application, the chapter's final two substantive sections consider some of the pragmatic issues that researchers confront as they implement surveys on the ground, such as costs, contracts, and managing survey teams. We explore the kinds of innovations, compromises, and work-arounds that can be necessary where methodologically ideal approaches prove infeasible. In particular, we address some of the complexities and pitfalls entailed in organizing people to make surveys happen. Of course, all of the

data-collection techniques discussed in this book can be undertaken either alone or in some form of collaboration. Survey research, however, tends to be particularly collaborative in nature. This may simply mean hiring research assistants to conduct interviews, or may entail more elaborate outsourcing to specialized organizations. Some well-known survey projects – the American National Election Studies and the World Values Survey, for example – involve collaboration on a truly massive scale, pooling the efforts of many investigators, coordinating through steering committees, and receiving multiple waves of funding over time. While many of the points we make could apply to survey research on any scale, our goal is to offer insights and strategies that will be useful for the kinds of projects that one PI or a small group of researchers might initiate.

## **Surveys' diversity and combining surveys with other data-collection techniques**

In survey research, information is gathered using systematic procedures from a population of people or other entities, or from a sample of a larger population.<sup>5</sup> Generally the sample size must be large enough to attain a reasonable degree of representativeness. Beyond these very basic similarities, surveys may be of many types, and may be combined with other data-collection techniques in several ways. This section considers these issues.

### **Diversity within survey research**

Typically the purpose of a survey is to generalize findings based on data gathered from a rigorously sampled subset to the broader population in ways that produce a calculable degree of error.<sup>6</sup> Often survey questions use pre-coded and standardized response categories, to be recorded in quantitative values for later analysis. Yet surveys sometimes do more than produce quantitative data, and the goal need not be generalizing to population characteristics. Surveys often include at least some open-ended questions

<sup>5</sup> This definition is adapted from Groves *et al.* (2009, 2). Narrower definitions exclude surveys of subjects other than individual people (Manheim, Rich, and Willnat 2002, 120). But, as discussed below, political scientists and other researchers often use surveys to explore other kinds of populations such as organizations or communities.

<sup>6</sup> For example, it might be calculated that, given sampling error, a survey's point estimate of the average age of a population is 95 percent likely to be within a range from 35.4 to 37.7.

either  
to be  
search  
ing to  
erican  
ple -  
many  
mul-  
could  
s and  
small  
  
from  
arger  
ain a  
rities,  
data-  
es.  
  
data  
ways  
pre-  
tative  
duce  
ation  
tions  
  
below,  
:  
verage

aimed at obtaining information from respondents in a more free-form manner. This material might later be assigned a set of quantitative or categorical codes, or it might remain in a qualitative format. Such qualitative material can serve many purposes. It can function as a check on closed-ended questions, allowing respondents to report things that do not fit into the survey's pre-established answer categories. It might capture information that is sufficiently complex to defy simple multiple-choice formats, such as the process through which a person became committed to a political cause, or the sequence of events in a legal dispute. The purpose might be to obtain an extensive list of items for follow-up study, such as names of influential people in a locality or web sites frequented by participants in a social movement. Nonetheless, most political scientists associate surveys with quantitative data, and producing such data is their comparative strength.

Over the years, political scientists have carried out a huge variety of surveys. With regard to the form of the survey, or its "mode," data may be gathered through traditional face-to-face interviews, or telephone interviews. Surveys may also involve mailed questionnaires, or may be email- or web-based. In all cases, researchers can contract out much of the work involved in a survey to other parties: survey research can thus be done without the PI ever leaving his or her home institution. In this chapter, naturally, we focus on projects that actually take researchers into the field.

In terms of subject matter, political attitudes, voting behavior, and civic engagement have long been core topics in survey research on the United States and other countries. But researchers in our discipline also explore many other topics through surveys, for instance, tolerance, interpersonal networks, and opinions on international relations. Scholars survey the opinions of political elites as well as the public. And far from confining themselves to individual human beings as subjects, they survey groups of people, firms and other formal organizations, and even places. For example, the maintenance of streets and the cleanliness of parks in small Midwestern towns have been surveyed, as one proxy among others for social capital (Rice 2001).

### Integrating survey research with other dimensions of a project

As examples mentioned throughout the chapter illustrate, researchers can bring information, familiarity, and social relationships from other phases of the research project and the field research endeavor into the survey process. Here we emphasize ways to integrate survey research with other forms of

data collection. Such integration is, in fact, overwhelmingly common in the field research projects reported in the FRPS survey of political scientists. Of the 410 reported field research projects that employed survey research, only 11 (less than 3 percent) involved no other data-collection techniques. Of those 410 projects, 89 percent also used interviews; 55 percent also involved archival research; 42 percent also featured ethnography or participant observation; and 25 percent also drew on focus groups.<sup>7</sup> Projects using surveys also used an average of nearly six (5.7) other techniques. And with regard to analytic techniques, in 85 percent of projects that used surveys to collect data, either qualitative or interpretive approaches or both were among the approaches to analysis that were employed. This highlights the fact that quantitative research often goes hand-in-hand with other forms of inquiry in field-based projects.

More than forty years ago, sociologist Sam D. Sieber published an essay discussing numerous ways in which surveys and fieldwork (as he understood it) could complement and inform one another.<sup>8</sup> Seeking to reconcile “two methodological subcultures” with a history of mutual antagonism, he suggested that each approach could guide the other, with survey results suggesting places for in-depth study, for instance. He also posited that findings generated through one method could be checked, validated, and interpreted by means of the other. Interviews and observation might provide a theoretical structure and hypotheses to be tested with large-*n* data, which in turn might correct for unconscious biases coloring the qualitative work, such as a tendency to place excessive weight on the views of elites, and the holistic fallacy (the “tendency on the part of field observers to perceive all aspects of a social situation as congruent”).<sup>9</sup> Qualitative research, Sieber suggested, could provide case studies to bring statistical relationships to life, while surveys could demonstrate the general applicability of such individual cases (Sieber 1973). His examples centered on sociological studies of schools, unions, and the like, but the arguments apply to a broad set of subjects.

Important epistemological contrasts notwithstanding, political scientists frequently synthesize surveys with more qualitative approaches to data

<sup>7</sup> The questionnaire asked respondents to indicate only data-collection techniques that their project made “significant use of.”

<sup>8</sup> Sieber defined fieldwork as “participant observation, informant interviewing and use of available records to supplement these techniques in a particular setting” (1973, 1335n2).

<sup>9</sup> As an example of a holistic fallacy corrected by survey results, Sieber discusses a project in which he observed staff members in one large and one small school district. Because staff in the small district treated one another with less formality, he initially inferred that their morale was also higher, which turned out not to be so.

collection. For instance, in a reflective methodological essay about one of her projects in China, Gallagher describes in detail how a set of fifty semi-structured, in-depth interviews with legal aid recipients helped reshape her approach to the large-scale survey that she later conducted, in particular changing the way she thought about key concepts. For example, she realized that the notion of “legal consciousness” contained complexities in the post-socialist context of Shanghai that were not found in the United States context where it had previously been theorized – complexities that could not necessarily be captured in a single linear scale. “The effect on the survey and how I thought about the survey was profound,” she writes, further advocating that qualitative interviews be seen as a crucial part of the research design process for survey projects (Gallagher 2013, 185). A scholar interviewed for this book related how an open-ended process of observation and interviewing was used to refine conceptual measures and build survey questionnaires: “I spend several weeks . . . soaking and poking in the communities that I intend to study. That gives me a much better idea of how the variables of interest should be conceptualized, and how the questionnaire should be designed.”<sup>10</sup>

On other occasions, multiple elements of an extensive and immersive program of field research can inform the design of a survey questionnaire. For example, during 20 months of research in Japan on siting decisions for “public goods” (airports, nuclear power plants, waste facilities, and incinerators), Daniel Aldrich employed multiple data-collection techniques, including conducting archival research and many in-depth interviews, carrying out extensive content analysis of newspaper articles, and building a database of localities and site choices for noxious facilities. Drawing on what he was learning from these other techniques, he designed a survey inquiring about things like compensation, local politics, and associations. He ultimately distributed the survey to more than 350 individuals in local governments and offices of the national infrastructure ministry. Not just knowing how to phrase the questions, but knowing whom to ask and what to ask about, required knowledge gained through the other activities (Aldrich 2008).

Another common practice is to *follow up* on survey findings by engaging in more in-depth interviewing of survey respondents or others from the target population, seeking to explore key concepts or causal processes, or bolster and explicate survey results. Indeed, the previously cited *Civic culture* employed such a strategy, re-interviewing 539 respondents from the national

<sup>10</sup> Interview, LM-10, September 18, 2012.

cross-section samples (about 10 percent) in order to obtain their “political life history.” The authors composed vignettes based around the life stories and world-views expressed in some of these interviews and compiled them in a long chapter illustrating what they depicted as the political cultures of the five countries under study. This approach highlights the rhetorical impact of such personal accounts (Almond and Verba 1963, 46, 402–469).

In another example, Strolovitch commissioned a telephone survey of leaders and staff at 286 advocacy organizations, then followed up through semi-structured interviews with some of the survey respondents, as well as interviewing others. The interviews, she writes, “supplement the survey data by providing a window into the nuances of how, why, and in what context organization officers make the decisions that they do about how to allocate resources” (2007, 240–248, at 247). In Read’s work on state-fostered neighborhood organizations in Beijing and Taipei, popular opinion surveys, which found a surprising degree of approval of state-backed groups, helped to frame and guide a series of interviews that explored the perspectives of city residents in more detail (Read 2012). In other instances, analyzing survey data may reveal a particularly puzzling distribution of responses for one question or by one subset of the respondents. This may lead to more focused follow-up interviews with a targeted subset of the survey sample to try to understand such “questionable responses,” as one investigator put it.<sup>11</sup>

One recent book exploring the critical role of the early Iowa caucuses in the US presidential nominating process demonstrates how surveys can be integrated with qualitative knowledge. A field survey was administered in January 2008 at the actual caucus sites of each political party in 1,784 Iowa precincts in 99 counties; one participant in each was given a written questionnaire before the beginning of the meeting, resulting in 2,611 completed responses (Redlawsk *et al.* 2011, 39, 88). The University of Iowa faculty who fielded the survey juxtaposed data it provided with information gathered through a more conventional telephone survey of caucus attenders, and data from other polls. They also integrated their own experience-based knowledge of the caucus process: the first author, for instance, had served five times as chair of his precinct caucus, and also as chair of a county-level party organization in Iowa. Through combining data from multiple sources, the study creates a rich portrait of citizen participation in this important event and builds an argument about the importance of grassroots-based caucuses in the election process.

<sup>11</sup> Interview, LM-10, September 18, 2012.

To offer a final possibility, other data-collection techniques can also be used to develop contextual variables that are subsequently linked to individual-level survey data. For example, Melani Cammett used GIS to create a map of the area where she planned to conduct a survey in Lebanon. She located in space the level and types of health care infrastructure in the sampled community. Later, she included in her statistical analysis a variable that indicated the distance to the nearest health center from the respondent's home.<sup>12</sup> In short, as these examples demonstrate, qualitative work can help create better surveys, and survey data can raise questions that can be pursued in greater depth through deeper qualitative follow-up.

## Why do a survey? Considerations and contributions to theory

Some field researchers set out from the very beginning of a project with a firm intention to conduct a survey, whether it is a central part of their research design or a peripheral component. Others may arrive in the field uncertain of what the possibilities are for doing a survey, particularly if the setting is an unfamiliar one. We highlight several factors to consider when assessing the pros and cons of including a survey component in a research project. We then discuss the broader issue of what analytic tasks scholars can accomplish with a survey – that is, how surveys contribute to the building of theory in political science.

### Factors to consider when contemplating a survey

A primary consideration when scholars are thinking of carrying out a survey, of course, must be the fit with the design and goals of the research. Survey research – alone or in conjunction with other forms of data collection – must show the promise of helping scholars to accomplish key analytic tasks such as descriptive inference or hypothesis testing (as we discuss in the next section), and bringing the overall project a considerable distance toward answering the questions it asks. More broadly, survey work only makes sense if it meshes with the researcher's skills, tastes, and epistemological orientation.

Before proceeding, it is sensible to determine whether existing survey datasets might suffice, making it unnecessary to start from scratch. As a leading survey expert cautions, “Sponsoring a special-purpose survey data

<sup>12</sup> Interview, LM-9, August 30, 2012.

collection is a rather expensive solution to an information problem" (Fowler 2009, 3). The search for extant datasets can, in fact, constitute a non-trivial component of fieldwork all of its own. In some settings, academic data repositories will provide an obvious first place to check<sup>13</sup> – and, convenient though they are, the process of sifting through their catalogs and poring over codebooks may take considerable energy. Even if such sources prove fruitless, that does not mean that no one has done the kind of research in question. Individual academics, private research firms and pollsters, and government agencies all may have conducted or commissioned surveys. In many cases, making contacts and asking around are the only ways to find data that are sitting in a drawer or on a computer, possibly available for the asking. It is this type of opacity and uncertainty that sharing data and research transparency – the latter of which we suggest is a principle of good field research – have the potential to mitigate. Yet as a practical matter, some datasets will always be withheld from publicly accessible archives, so there is no substitute for making private inquiries.

Yet even such extensive investigation may not turn up survey data relevant for a scholar's particular research questions, field sites, or time periods. Existing datasets may not focus on the precise questions of interest or operationalize and measure concepts in the way that the researcher finds most compelling. Aggregate datasets are often biased toward more stable countries and regions where survey research firms can operate safely. For example, with the exception of Zimbabwe, the Afrobarometer Data Project has historically included countries that have at least moderate levels of state capacity and are more democratic and economically open. In countries and regions still troubled by war or only recently emerging from conflict, few or no survey data may exist. At other times, the groups a scholar wishes to study may have been neglected in previous survey research. One researcher we interviewed explained how she became one of the very first to conduct a survey with women farmers in a post-conflict setting. Almost no systematic data had been produced in this region and with this population during a long civil war, where many of the roads leading to the location were either

<sup>13</sup> Perhaps three of the most well-known archives are those available through the University of Michigan's Interuniversity Consortium for Political and Social Research (ICPSR), the University of Connecticut's Roper Center, and Harvard University's Dataverse. The Council of European Social Science Data Archives (CESSDA) is an umbrella organization that maintains a catalogue of data repositories across Europe. Institutions around the world host data repositories; the University of Amsterdam's Social Science Information System (Sociosite) keeps an extensive list, sorted by country.

destroyed or covered with mines.<sup>14</sup> Conducting one's own survey hence allows scholars maximal freedom to pose exactly the questions they wish to ask to the right people, in the right place, at the right time.

The costs of a survey, in terms of both time and money, bear careful consideration. The factor of time is easily underestimated. Even with what appear to be straightforward and small-scale surveys, the many steps in the process – at minimum, defining the scope of the project, writing questions, designing the questionnaire, hiring and working with any partners involved, pre-testing, sampling, and then collecting, cleaning, processing, and analyzing data – almost always take more time than anticipated. Financial costs vary widely with such factors as the survey's location, mode (for instance, in-person interviews vs. mail or telephone), partners, sample size, degree of methodological rigor, and fit with the established templates of previous projects. Depending on the setting, researchers, even at the graduate student level, should not assume that the monetary cost is beyond reach. For modest surveys, it may be possible for the researcher to do much or all of the work himself to save money. Alternatively, survey organizations can often tailor their work to fit a fixed budget, and may be able to include a small set of questions from one's project on a survey planned for another client. Moreover, a well-designed proposal showing precisely how the data will pay off might be just the thing to attract the favor of a funding organization.

Finally, particular settings may pose their own special challenges to fielding a survey. For example, authoritarian states like China impose administrative constraints on survey research related to social and political issues, and investigators from outside of mainland China are subjected to extensive oversight and special restrictions. The Chinese National Bureau of Statistics and its local counterparts assert jurisdiction over surveys, and often other government agencies request payment as well. Well-connected partner institutions are thus required to carry out survey research in this setting. While noting this imposing regulatory regime, Manion pointed out that still no fewer than thirty-two survey projects in China from 1986 through mid-2008 had resulted in peer-reviewed articles, chapters, and books by political scientists writing in English (2010, 186–187).<sup>15</sup> An overview of research done on public opinion in the Arab world pointed out that, although for decades such studies had been subject to stringent limitations, political change and other factors created an opening for new projects (Tessler and Jamal 2006).

<sup>14</sup> Interview, LM-17, September 11, 2012.

<sup>15</sup> This tally included only publications in academic venues outside of China.

In short, political obstacles are among the concerns that must be evaluated in deciding whether or not to pursue a survey, but, like other potential stumbling blocks, they vary in seriousness and can often be overcome. We discuss some related challenges, and strategies for addressing them, in subsequent sections of the chapter.

### **Surveys' analytical contributions**

As we noted in this chapter's introduction, data garnered through survey research may contribute to accomplishing a variety of analytic tasks, in four general ways. They can enrich descriptive inference, for instance, by allowing researchers to examine the distribution of variables across a population or compare subgroups. They can allow for testing hypotheses and, in some circumstances, exploring causal relationships. They can help the investigator refine aspects of the project to be explored in later stages, for instance by revealing which cases deserve special scrutiny or what new hypotheses deserve to be tested. Finally, and less commonly, they can also aid with conceptualization. We consider each in turn below.

#### **Enriching descriptive inference**

Surveys offer particular ways of enriching a scholar's descriptive understanding of phenomena under study, sometimes enabling inferences that are impossible or difficult to draw through other data-collection techniques. Surveys may be able to provide information about the political attitudes, behaviors, and experiences of a greater number and variety of respondents than other methods. Most often using a random sampling technique to obtain a representative sample, surveys can help researchers to examine the distribution of key variables in the broader population. This facilitates comparison across subgroups and helps identify patterns, trends, and anomalies. Survey research may capture a greater range of citizen viewpoints than can be obtained from subjects purposefully selected for individual or focus group interviews. In sum, survey research makes several important contributions to descriptive inference, a set of tasks whose value is sometimes discounted by political scientists but that is often an essential part of a field research project.

#### **Testing hypotheses**

The second way that surveys may contribute to theory is by enabling the testing of hypotheses. For many researchers, a primary rationale for going to the trouble of conducting a survey is to obtain a dataset allowing for statistical testing of important propositions about an underlying population,

such as whether voting turnout rates are higher among voters in a particular age or ethnic category, or whether citizens who are members of a ruling party are more likely than others to receive welfare benefits.

Depending on the exact proposition at issue, and the kinds of assumptions one can make, it can be difficult to assess causal hypotheses using standard cross-sectional surveys. Correlations between observed variables do not necessarily establish whether X caused Y, Y caused X, or both were caused by a third, perhaps unobserved, factor. Experimental designs (discussed in the next chapter) are one response to this problem. One subset of experimental research, survey experiments, involves manipulating the wording or ordering of questions in a survey and assessing the effect on respondents' answers. This approach makes sense where a posited causal factor is amenable to manipulation and random assignment within a survey, such as in a study of how the language in which a policy is described affects support for that policy.<sup>16</sup> Survey researchers take other approaches as well. The use of panel surveys is one approach, such as Baker, Ames, and Renno's study of preference change among voters in Brazil's 2002 presidential election, which successively surveyed the same voters at three points during the campaign (Baker, Ames, and Renno 2006). Brady discusses other approaches, such as quasi-experimental designs (2000, 52–53).

### Selecting cases and generating hypotheses

Surveys can also be used as a source of information that shapes the design of later components of a field research project. Though it was not written with survey research in mind, Lieberman's discussion of "nested analysis" suggests logics for selecting cases for in-depth study on the basis of findings from surveys.<sup>17</sup> Large-*n* analysis might, for example, generate a regression line positing a relationship between two concepts. Follow-up case study work

<sup>16</sup> See Mutz (2011) for a detailed treatment of population-based survey experiments. Gaines, Kuklinski, and Quirk (2007) offer a critical overview of the subject.

<sup>17</sup> In Lieberman's article, the units of analysis are specified to be such things as nation-states, provinces, or institutions. He writes that "for most analyses of individual behaviors or attitudes, for which the 'large-N' component of the data is contained in a survey, I would not expect this approach to be feasible, because scholars are unlikely to be able to conduct further in-depth research with the original respondents. Moreover, the prospect of explaining the exceptional nature of a particular individual is unlikely to be of intrinsic interest in the way scholars are likely to be interested in the particularities of larger social units, such as national states" (2005, 436n2). Follow-up with individual survey respondents can certainly be possible, though, and may be valuable for obtaining evidence about what explains observed correlational patterns. Also, surveys are not always of individuals; they sometimes have supra-individual entities (neighborhoods, villages, firms, institutions) as their primary unit of analysis, or individual-level responses may be aggregated into such larger units. Thus, Lieberman's logic can indeed apply to survey work.

could focus on representative cases on the line, for model-testing analysis, or could include off-the-line cases for more preliminary model-building work (Lieberman 2005, 440–446). Another way in which surveys can drive follow-up research is that open-ended survey questions can be used to produce potential explanations that can then be explored as alternative hypotheses. For example, respondents might be asked their opinion of a politician, policy, or law in a closed-ended question with answer categories amenable to quantitative analysis, such as a feeling thermometer. They might then also be asked *why* they indicated a positive or a negative opinion, and these answers could be explored in later in-depth interviews, or else coded and analyzed in quantitative fashion.<sup>18</sup>

### Developing concepts

Surveys can also contribute to conceptual development. Survey researchers have to decide, in advance, precisely how they will word each question, structure their instruments, and code respondents' answers. This encourages careful consideration of the underlying concepts. Brady notes that survey researchers spend "extraordinary amounts of energy thinking about how to devise measures of concepts by asking questions," and suggests that controversies concerning measurement in survey question wordings have had the effect of sharpening our grasp of underlying concepts. He cites, as an example, the problem of how to assess a respondent's political tolerance. This has been measured through questions about respondents' willingness to acknowledge the civil liberties of groups they dislike, raising the problem of what kind of group makes an appropriate referent from which to gauge tolerance (Brady 2000, 51). In that example, the concept in question has evolved over the course of decades of study, but a single survey can also offer possibilities for probing the meaning of a concept. For instance, a researcher might include multiple, differing measures of a single concept, or ask respondents to rank answers related to a concept in order of importance or priority.

---

### Preparing for surveys in the field

---

As with other data-collection techniques, administering a survey in the field frequently involves making fundamental research design decisions after arriving in the field site. While a scholar will likely dedicate extensive thought

<sup>18</sup> Interview, BR-3, August 6, 2012.

to designing the survey prior to leaving her home institution, such basics as the structure of the survey instrument, sampling technique, sample size, the mode of the survey, and question selection, ordering, and wording may all be affected if not dictated by the local context. Making decisions about all such matters, and many other details, is part of why preparing to field a survey can and should take substantial amounts of time. One interviewee passionately advocated allowing enough time at the beginning of the project to figure out how things work on the ground and what questions are important, rather than bringing in a preformatted questionnaire, which (as the scholar put it) might reduce our collective knowledge instead of adding to it.<sup>19</sup>

Refining and pre-testing a survey present many opportunities for correcting omissions, catching mistakes, and ensuring that the survey meets the researcher's needs. But surveys may not be open to *as much* ongoing refinement as the other data-collection techniques discussed so far in this book allow. One might do months of ethnography, for example, and in the last week begin probing a new set of questions or issues, with no great harm done. With a survey, the investigator reaches a more definite point of no return. Once the survey is launched or goes "into the field" for primary data collection, changes in overall design or even in tiny details of question wording become extremely costly. Such modifications may mean that data acquired in the second iteration of the survey are incompatible with data gained through the first iteration, and thus of limited or no value.

In short, the iterative process of ongoing research design – which we find to be widespread in field research and whose merits we emphasize – bumps up against some limits in a survey project. Consequently, scholars who contemplate carrying out a survey (or a field experiment, as discussed in the next chapter) need to anticipate their data needs carefully in the early phases of the research process. Indeed, it is in part because of the highly structured – even inflexible – nature of survey research that surveys are typically fielded later in a research project, once many key analytic tasks have already been accomplished, often by employing other data-collection techniques.

The rest of this section discusses several elements of survey planning that scholars will need to carry out in the early stages of their project – although, again, some tasks will need to be completed after arrival in the field. We begin with research partnerships, considering the options of subcontracting a survey firm versus training one's own research team. We then consider designing questionnaires, sampling, and pre-testing in the field.

<sup>19</sup> Interview, LM-10, September 18, 2012.

### **Subcontracting a survey or training a research team**

Having opted to pursue a survey, one of the fundamental questions for the researcher is whether to contract out the labor involved to a professional survey organization, to do it herself or with the help of research assistants, or to employ some mixture of these two strategies. The choice may be dictated by local circumstances and the availability of reputable firms and competent assistants, but, in generic terms, each approach entails fairly predictable sets of advantages and disadvantages. Beyond a relatively small sample size, just about any type of survey requires the work of multiple people. Organizations dedicated to this kind of research have staff and equipment on hand, and using their resources can be more efficient than trying to do all the legwork oneself. As well, private firms or academic survey centers have experience in key realms such as the creation of workable sampling frames and the wording of questionnaires, and can provide insight on matters like expected response rates for different approaches.

Conversely, there are also advantages to running the survey on one's own. In some parts of the world, reliable research organizations may simply not exist, or may be prohibitively expensive to contract. Also, the investigator may want the survey to use methods or include features that lie out of the norm for firms whose bread-and-butter work is market research rather than academic inquiry. Finally, self-reliance provides maximum control to the scholar over the integrity of the process – and no one cares about the quality of the data like the PI herself.

In exploring possible collaborations with survey firms and organizations, it is critical to shop around. By eliciting bids from multiple vendors and, when possible, visiting them in person, scholars can get a good sense of how the firms operate. Researchers might also get recommendations from academics who have worked with the firms. Firms with a track record of doing the kind of work that is being planned, whose staff show a substantive interest in the project or the topic of inquiry, and that seem to have a strong commitment to quality might be particularly good partners. Indeed, it is important, even in these early stages, for the PI to set the tone that quality is of paramount importance and that all work relating to the project needs to be done conscientiously. Once scholars begin to negotiate with one or more firms, they should write up the critical project specifications, from sampling to the handling of non-responses to payment, in a scope-of-work document to ensure that everyone understands the expectations for the project.

Investigators with experience working with survey organizations emphasize the importance of taking an assertively hands-on approach to the partnership throughout the process. If possible, the researcher should be present at key stages, such as when enumerators are being trained, and should quickly address anything that deviates from his expectations. One political scientist recounted an incident in which a firm she had hired delivered data to her that required many hours of cleaning and fixing: “The lesson from that was that I wasn’t doing enough homework all along; I needed to be on top of them every step of the way. You should review what’s coming in and not release payments until you are satisfied.”<sup>20</sup> (This scholar decided to train her own team in future projects.)

While PIs need to stay closely involved with their project regardless of the quality of their partner firm, it is important to highlight that survey research organizations with proven track records can contribute knowledge and experience to a project. A researcher who conducted surveys in multiple countries overseas found staff in reputable survey firms to be “experts in and of themselves,” providing valuable perspective on sampling techniques, language issues, and how questions on particular topics had fared in previous polls.<sup>21</sup> It can also be possible to save money and maintain a higher degree of personal control by contracting a survey outfit for only certain services and not others. For instance, a researcher could hire a firm’s enumerators to conduct survey interviews, while doing sampling and data entry herself (or with a team of research assistants). No matter how much or how little of the work is outsourced, researchers should stay involved and keep on top of what is being done.

Forgoing survey organizations and training one’s own enumerators brings with it a distinct set of advantages and disadvantages. Doing so can be the most economical option, and it leaves day-to-day responsibility and oversight in the researcher’s own hands, maximizing her ability to ensure that correct procedures are followed, errors are corrected, and the resulting data meet the desired quality standards. One political scientist explained how he had administered a survey in two overseas cities himself, managing a set of supervisors and building a team of dozens of college students as interviewers. Because the interviewers had no prior experience, they required considerable training and varied in quality. The intense and protracted work, and countless details, including resolving squabbles within the team, created

<sup>20</sup> Interview, LM-9, August 30, 2012.

<sup>21</sup> Interview, BR-11, August 22, 2012.

headaches: “I was miserable the entire time.”<sup>22</sup> The money the scholar saved by not hiring a professional firm, however, allowed him to draw a much larger sample. And, of course, working as part of an international research team can be an enriching experience for students with regard to both methodological knowledge gained, and in-country and international connections made. Another experienced researcher suggested that it is possible to economize by employing local graduate students as enumerators and compensating them by letting them contribute questions designed to advance their own projects.<sup>23</sup>

One comparativist arrived in-country and hastily hired an enumerator in hopes of getting a fast start on a survey. Having been paid half his compensation in advance, the enumerator became frustrated with the challenging work and “ended up running away with my money.” Learning from this experience, the researcher vetted assistants more carefully, recruiting them through universities and also through institutions with strong ties to the communities that were the targets of the surveys. Paid on a daily basis and working side-by-side with the PI, later enumerators not only accomplished the primary goal of completing the survey interviews but also provided perspective on their interactions with respondents.<sup>24</sup>

Several of our interviewees were convinced of the merits of organizing and managing the survey process themselves – both with regard to what *they* learned through the process, and in terms of the quality of the result. A scholar of public policy declared resolutely that the “huge investment” in training his own research team was recouped in his confidence in the quality of the data: “Farming this stuff out would have resulted in junk!”<sup>25</sup> A comparativist commented:

In the end, I was glad I didn’t work with a survey firm. If I had, I would have outsourced much of what I did to them. I wouldn’t have had as much influence over the sampling, wouldn’t have been able to have my hands on everything. In the end, I did everything myself, including stapling the questionnaires . . . basically I never slept for weeks. I really saw it from the inside. In terms of learning how to do a survey, I think it was incredible training.<sup>26</sup>

Whether one is working through a survey organization or not, the training and management of the staff who will conduct the interviews or otherwise

<sup>22</sup> Interview, BR-3, August 6, 2012.

<sup>23</sup> Interview, BR-15, October 25, 2013.

<sup>24</sup> Interview, BR-6, August 14, 2012. Bonilla also emphasized the value of “the interviewing of interviewers” (1964, 149).

<sup>25</sup> Interview, LM-10, September 18, 2012.

<sup>26</sup> Interview, BR-14, October 24, 2013.

stand on the front lines of data collection are paramount. Survey researchers of all stripes agree on this. If interviewers and enumerators shirk or deviate from their responsibilities, or simply fail to put energy into obtaining the cooperation of the subjects of the survey, quality will suffer. (A different problem is putting too much energy into this task, such that interview requests and questions border on the coercive, an experience MacLean had with one of her enumerators early on.) In short, because every project and its emphases, imperatives, and procedures are different, even staff members with previous experience in survey research should receive careful training. Training should be wide-ranging, addressing each step of the survey process including approaching respondents, eliciting informed consent, interacting with respondents as questions are asked, capturing responses and additional information about the interaction, wrapping up the session, potential follow-up, and much more.

Likewise, once the survey is being fielded, a considerable degree of monitoring and oversight is always required in order to ensure that the project is being carried out as the researcher envisions, and to identify and resolve problems. Simultaneously, and just as importantly, scholars should do all they can to help and coach those whose help they have enlisted. Researchers should make themselves available to answer patiently any questions that arise. It can be useful to have periodic meetings with all team members – to address issues and resolve questions, to keep the whole team updated on progress, and to help to build team spirit. Some researchers stress other strategies for reinforcing team members' commitment to the integrity of the research process. Comparativist Lily Tsai personally trained and supervised three teams of ten to twenty students each for her village-level surveys of public goods provision in rural China. She describes traveling with them via bus and train, emphasizing the intellectual and social value of the project in order to increase their commitment, and reinforcing bonds of mutual obligation with them through practices "from advising them on their theses to staying in the same accommodations to hand-washing their laundry when they were busy with survey administration" (Tsai 2010, 258). While not everyone will go to such lengths, most would echo the idea of trying to build a cadre of dedicated research assistants and working with them directly and carefully.

### **Questionnaire design in the field**

Specialized methodological monographs and general textbooks provide critical guidance on the general methodological issues involved in the writing

of a survey (Converse and Presser 1986; Fowler 1995, Krosnick and Presser 2010). But crafting questions that address the goals of the research project, that are cast in appropriate terms, that make sense to respondents, and that do not provoke unwanted reactions requires considerable local knowledge as well as proficiency with the general principles of survey research. Indeed, as the examples of combining surveys with other data-collection techniques offered earlier in the chapter illustrated, this is an area that can be significantly informed by data gathered through other aspects of the field research project.

The need to adapt a research design into locally meaningful terms that can be operationalized in a survey questionnaire starts at a high level of generality. What are the underlying concepts that the project aims to assess? How are those concepts manifested in the field environment and in the subjective world of the people who will be enumerating or responding to the survey? Of course, researchers should avoid allowing social science jargon or unfamiliar terms to slip into the questionnaire. Whether in the native tongue or a different language, the words of the survey instrument should probably not be the words used in the project's grant applications or prospectus. Yet even when researchers have successfully accomplished that aspect of "translating" their research design into their questionnaire, they may discover, once they arrive in the field, that the questions they planned to ask – or the particular way they planned to ask them – do not fit the local setting well, and must be rethought. These issues, of course, derive from the fundamental tension in social science between general concepts and the endless particularity and variation of the empirical universe. One example of a controversy of this kind revolving around survey measurement is found in the debate over how "civil society" should be conceptualized and how it should be measured in different contexts.<sup>27</sup>

The practical implication is that researchers need to find appropriate and meaningful ways to connect their theoretical concerns with local realities. The ways to do this are many, of course. Scholars might scour publications based on studies pertaining to the concepts in question and carried out in the setting in which they will field their survey; talk with local colleagues about concepts and how best to express them; and become familiar with settings like the ones in which they will field their survey. For instance, one scholar noted that it was not clear what the concept of social capital was or how it

<sup>27</sup> See the symposium on conceptualization and measurement in the *Journal of Civil Society*, vol. 1, no. 3, December 2005. Related controversies are addressed in papers concerning the World Values Survey, [www.worldvaluessurvey.org/index\\_paperseries](http://www.worldvaluessurvey.org/index_paperseries).

would manifest in a context that differed from Putnam's Italy or Coleman's United States. Nearly three months of intensive qualitative interviews and ethnographic observation were needed to develop an appropriate scale and instrument.<sup>28</sup>

At the level of question wording and (if necessary) the translation of terminology into local languages, complexities abound. Even with seemingly straightforward items of the kind that it might seem could be asked almost anywhere – questions about respondents' dates of birth, marital status, or line of work, for instance – it is important to seek guidance on phrasings in cultures that are not the researcher's own. In Taiwan, to give just one example, it is customary to express years, including birth years, in the numbering system of the Republic of China calendar, counting from the 1911 revolution. Questions about household income are sensitive in many contexts, Western and non-Western, and often require careful wording and the use of approximate ranges. Indeed, one researcher found that poverty was such a stigmatized term that enumerators had to first discuss it "hypothetically" and then circle around to speak of poverty in terms of visible concrete items in the household.<sup>29</sup> Importing question wordings from surveys in other languages almost always presents complications. The task is as complex as any problem of translation, but more acute than most because respondents should not have to strain to make sense of the wording; it must come across as natural to them.

The need for locally appropriate language is even more obvious when it comes to concepts that are specific to the locality: religions, political parties, notions of approval or disapproval and trust or distrust, government agencies, and the like. Professionals and academics with experience polling in the same context can help scholars perfect their questionnaires, and suggest prior surveys that have been fielded in the same context, from which wording can be adopted. Questions can also be tried out on local research assistants, friends, or others in the field.

Although crafting locally suitable questions is of paramount importance, the value of maintaining compatibility with concepts and wordings that are already in wide use in the discipline also deserves emphasis. Many key political variables – party identification, tolerance, political support, etc. – have long track records of operationalization in surveys around the world. Straying from previously accepted wordings can make it more difficult for a researcher to compare her findings with those of other surveys, or to

<sup>28</sup> Interview, LM-10, September 18, 2012.

<sup>29</sup> Interview, LM-10, September 18, 2012.

speak to longstanding debates. For this reason, it may be worthwhile to triangulate by including multiple measures of important concepts, for instance one in more “localized” terms and another using phrasing that more closely follows standard international practices.

## **Sampling**

Ideally, sampling entails defining a full population and then randomly selecting a subset for study in such a way that each unit (e.g., individual person or household) within the broader population has an equal chance of being included. Obstacles in the field can complicate sampling, however, necessitating creativity on the part of researchers. A central problem is that for many of the populations that interest political scientists – grassroots NGOs, or informal worker organizations, or citizens who have appeared before local courts – the full population is difficult to define and no list of that population exists. Moreover, even when a list of the target population does exist, it can be inadequate or systematically biased in ways that complicate sampling. For instance, in the United States, sampling from a frame of possible land-line telephone numbers via random digit dialing was once widely considered a good way to survey the general population for many purposes. Since the mid-1990s, however, the practice of individuals having no land-line but only a cell phone has presented an increasing challenge for surveyors and pollsters.<sup>30</sup>

Political scientists have employed a wide range of sampling strategies when conducting surveys in the field, from simple designs to more sophisticated ones employing stratification, clustering, and multiple stages. Interval sampling (such as selecting every seventh individual or unit from an arbitrary starting point) or random-walk sampling (starting at a given point and selecting the individual or unit found after proceeding along a random route) can be effective strategies in cases where no complete population list is available. Sometimes, and for certain purposes, approaches like quota sampling, convenience sampling, anchor sampling, and snowball sampling have to suffice; political scientists and others have also developed methods for improving

<sup>30</sup> This is addressed at length in a task force report by the American Association for Public Opinion Research, “New considerations for survey researchers when planning and conducting RDD telephone surveys in the U.S. with respondents reached via cell phone numbers,” 2010, accessed February 23, 2014 at [www.aapor.org/Cell\\_Phone\\_Task\\_Force\\_Report.htm](http://www.aapor.org/Cell_Phone_Task_Force_Report.htm). See also Kempf and Remington (2007) and Link *et al.* (2007).

the samples that such non-probability techniques provide.<sup>31</sup> In all types of sampling, the scholar should attempt to discover and document as much as possible about the characteristics of the population as well as subgroups within it. This information can be indispensable at later stages when considering whether and how to weight the data. All sampling strategies entail tradeoffs that researchers must carefully assess, bearing in mind their intellectual aims, local conditions, and the kinds of audiences (e.g., particular journals and their reviewers, dissertation committee members) that will be assessing the choice of sampling technique. We offer some thoughts on these issues in Chapter 3.

Each field setting presents its own kinds of sampling challenges. In most East Asian countries, for example, all households are supposed to be registered in a set of records maintained by the government, which document who lives at what address and include basic demographic information. These records are an attractive resource for survey researchers, as they hold the promise of providing a complete list of the population from which to sample. Indeed, their use is ubiquitous: many pieces of survey research, in China for example, have been based on samples drawn from such household registry records. But these lists pose problems as well. For instance, the Chinese population is divided between those who possess urban household registrations (or *hukou*) and those with only rural documentation. Since the end of the Mao era in the 1970s, rural residents have migrated to cities in immense numbers. Though they sometimes obtain temporary residence permits, they are not included in the regular household records, thus creating a substantial gap between who actually lives in the city and who is present on paper. Even among permanent urban residents, these household lists are only an imperfect guide to who actually lives where, as it is common to move without updating one's paperwork, or to maintain more than one residence. Sampling from these lists can thus introduce systematic bias, leaving out rural migrants and other mobile members of the population.

This problem has given rise to some creative solutions. One approach is for researchers to dispense with the imperfect local system of official household rosters completely, instead making their own list of the residences in a neighborhood. An ambitious way of coping with the problem is the spatial sampling approach developed by Landry and Shen. In this approach,

<sup>31</sup> These ideas were raised during a manuscript review session held at Indiana University, December 2011. Alexandra Scacco's paper, "A snowball's chance in Nigeria: finding rioters using respondent-driven sampling," also discusses such techniques. Accessed October 5, 2012 at <https://files.nyu.edu/als8/public/research.htm>.

“township” is the secondary-level sampling unit, and “household” is the primary-level sampling unit. Once townships have been randomly selected, grid squares defined by lines of latitude and longitude are drawn within them and randomly sampled. Trained surveyors enumerate all the households within those squares, for further random sampling (Landry and Shen 2005; Landry 2010). The resource-intensive nature of this approach means that it is unlikely to supplant simpler methods. What is important to note, however, is that, even given its technical sophistication and potentially (if not practically) universal applicability, this approach was developed and implemented through deep knowledge about specific circumstances in China, from housing densities to the training of enumerators.

A few other innovative sampling methods that political scientists have devised in the absence of an adequate list of the target population bear mentioning. When Berry, Portney, and Thomson began their study of urban citizens’ participation in local politics, they sought to assemble a sample of civic leaders and other city elites (among other data sources) in each of five US cities. They included in the sample all city councilors and heads of relevant municipal agencies. In order to ensure the sample also included individuals from the more ephemeral population of grassroots civic activists, they compiled lists of names that had been suggested in responses to a broad mailing to local leaders of many kinds, and by people interviewed in the course of their fieldwork (Berry *et al.* 1993, 310–311). In research on small-scale, informal-sector businesses and the state in Peru, Roever needed to build a sampling frame of street vendors and microentrepreneurs. Because no list of such people existed, she created one from scratch. Over the course of ten days, her team of fifteen enumerators methodically canvassed a numbered set of city blocks, as well as the rows of stalls inside commercial centers, for each of two business districts in Lima (Roever 2005, 201–206). Scott Straus visited fifteen prisons in which perpetrators of the 1994 Rwandan genocide were held, obtaining and sampling from lists of inmates meeting particular criteria – those who had confessed and been convicted – a striking example of a sampling design that could hardly be conceived without detailed knowledge of the Rwandan justice system gleaned from fieldwork (Straus 2006, 97–103).

### **Pre-testing in the field**

With any form of survey research, carefully pre-testing the survey instrument in advance of formally fielding the survey is *de rigueur*. Inevitably, the initial draft of a questionnaire will need revising to improve question wording,

answer categories, comprehensibility, and more. Often many rounds of pre-testing and revision will be necessary, and researchers' schedules must take this into account. For example, even after multiple pre-tests and extensive enumerator training, one scholar sampled three extra villages to implement several complete dry runs of the survey before finalizing the questionnaire.<sup>32</sup> Practicing administration with members of the target population or role-playing proxies from among acquaintances or fellow researchers provides a shakedown of all relevant steps and procedures; the PI can thus identify problems and adjust and revise accordingly, helping him to avoid design-compromising changes further down the line. With regard to the survey instrument itself, pre-testing will show whether, for instance, the human subjects protocol is understandable to the target population, the questionnaire is too long or confusing, or additional transitional language is needed between questions. Focus groups or debriefing of respondents after the pre-test can also suggest adjustments to the protocol. Performing basic analysis on the data generated via the pre-test also provides an opportunity to catch important omissions, for instance, of control variables.

Particularly in the field context, the pre-testing phase of a project also provides an opportunity to bolster the training of research assistants or interviewers and, potentially, to strengthen their investment in the project. For instance, PIs can reinforce the importance of subjects' informed consent and the procedures the researcher has established for obtaining it through the pre-test process. That process also allows for discussing and explaining appropriate techniques for probing subjects for answers without leading them in any direction. Researchers can also review questionnaires as they come in to see what administration problems they may suggest (as several of our respondents suggested they should do throughout survey implementation), even if immediate data entry is not possible.<sup>33</sup> If the data from the pilot questionnaires *can* be entered and examined immediately, it may be possible to identify and remedy any remaining problems with the training of enumerators before the survey is formally implemented. Pre-testing also helps research assistants to become better at administering the survey and more confident in their ability to do so, and thus more effective. Further, enumerators who are encouraged to suggest revisions to enhance comprehension and flow may feel greater ownership of the questionnaire and the

<sup>32</sup> Interview, LM-10, September 18, 2012.

<sup>33</sup> Interviews, LM-9, August 30, 2012; LM-13, September 7, 2012.

project as a whole.<sup>34</sup> Finally, a real sense of team spirit can be generated as bonds form and excitement builds through the pre-test process – a camaraderie that can be cultivated through the rest of the project.

---

## Engaging with sites and respondents

---

Writing in the early 1960s, Bonilla took on the question of how well the model of public opinion surveys prevalent in the United States applied to research in developing areas. He noted that the encounter between survey interviewer and respondent, conventional and straightforward though it may seem, in fact depends on shared assumptions and norms of communication that can be very context-dependent:

A total stranger appears, usually unannounced, and demands admission to the home. He proceeds to extract information about the family relationships of all who occupy the dwelling and then seeks to isolate one specific individual for more extended interrogation. He insists that the ensuing dialogue adhere to a rigid and unfamiliar pattern, frequently giving exact instructions as to the form in which he wants replies and sternly discouraging departures from his prescriptions. Though the subject matter and the phrasing of questions may seem to the respondent argumentative, embarrassing, gratuitously aggressive, or even dangerous, the uninvited visitor proceeds impassively, taking little note of the interviewee's distress or exasperation, all the while refusing to reciprocate by revealing his own sentiments regarding the matters under discussion. (Bonilla 1964, 140)

While modern polling is anything but a novelty in some of the places Bonilla discussed (such as Rio de Janeiro), in many contexts survey interviews with strangers remain a far-fetched or hazardous idea. In developing their strategy to encourage participation in a survey, researchers should carefully consider how respondents might experience their interaction with an enumerator, and how the encounter can be made as smooth and comfortable as possible for them. Taking such steps is the ethical thing to do, and may ultimately allow the survey to generate better data. Developing an effective strategy is far easier, we contend, the better a scholar understands the context in which she is operating.

There are many reasons why members of the target population may be reluctant to agree to take a survey – or to answer particular survey questions. They may simply not wish to spend their time talking to an interviewer or

<sup>34</sup> Interview, LM-10, September 18, 2012.

filling out a questionnaire, or may be impatient from too many telephone marketing calls. In some places, fraudsters pose as pollsters, and criminals have been known to use fake surveys to identify vulnerable homes to rob (a reason to exercise restraint, in some contexts, in asking for identifying information). In others, the government or previous colonial rulers may have a history of sending agents to register land or to document revenue and assets for taxation purposes. Overall, the locality may be politically or culturally inhospitable to research. Finally, as the paragraph from Bonilla indicates, surveys from political scientists may include questions on topics that respondents are uncomfortable talking about. They may solicit opinions, or information about affiliations or actions, that respondents would not normally share with strangers, and whose revelation could create awkwardness or conflict with family members, neighbors, or co-workers, or get them into serious trouble.<sup>35</sup>

The problem of how to encourage respondents to answer questions on sensitive topics is one that has generated a substantial methodological literature. Yet any strategy adopted will need to be adapted to the field setting. For instance, a textbook suggests that “embedding one sensitive question (for example, an item on shoplifting) among other more sensitive items (an item on armed robbery) may help make the sensitive item of interest seem less threatening by comparison.”<sup>36</sup> About this, one interviewee working in a tense political environment overseas remarked: “Maybe this works in the U.S., but I thought this was terrible advice in my context.”<sup>37</sup> Instead, this researcher designed a separate questionnaire for sensitive questions that respondents filled out while the enumerator stood some distance away, and placed in an envelope for added privacy. In another example, Bleck had respondents take a survey on an iPad, allowing the respondent alone to hear sensitive questions. Further, some response codes obscured respondents’ actual response (e.g., “fish” stood for “protested often”), and in other instances respondents simply pressed a picture on the iPad to respond rather than answering audibly.

<sup>35</sup> The list of topics in political science that people taking surveys might find sensitive is long. To name just a few, respondents might be hesitant to discuss corruption, being offered money in exchange for their votes, participation in protests or opposition activities, or even their true opinions about the ruling party. Work concerning sensitive topics in surveys by political scientists and others includes Silver, Anderson, and Abramson (1986), Berinsky (1999), Tourangeau and Yan (2007), and Ocantes, de Jonge, and Nickerson (2013).

<sup>36</sup> This is found in multiple editions of Groves *et al.*, *Survey methodology* (e.g., 2009, 247–248).

<sup>37</sup> This interviewee (BR-14, October 24, 2013) felt that adding material that was even more sensitive to the survey question would have backfired, thoroughly alienating respondents.

In settings where access to the population under study and sensitivity of research questions pose major problems, some survey researchers employ a strategy of patiently building personal trust and understanding with respondents, akin to that discussed in the chapters on interviewing and site-intensive methods. Jamal's research on associations in Palestine provides an example (2007). This project involved a national survey (carried out by a survey organization), in-depth interviews, and a survey of 425 association members. In a co-authored article, Jamal discussed this latter component of the project, conveying the painstaking approach that she took and the conditions that called for it:

Face-to face interviews were necessary to ensure the participation of respondents. Sometimes, I would have to meet several times with respondents before they agreed to participate. I would often have coffee or tea with them beforehand. I invested a lot of time in casual conversations and in getting to know the potential respondents before we conducted the survey. I discussed my project in detail and explained how their participation was vital to the success of the project. (Tessler and Jamal 2006, 436)

General political tensions in Palestine, skepticism about her credentials and affiliation, and subjects' unfamiliarity with this kind of social science research all necessitated special trust-building efforts:

Further, because most of the population had very little experience in responding to surveys, they wanted to think through their responses very carefully. They often asked for follow-up explanations – which took more time. Furthermore, they sometimes would ask me to answer the question first; they wanted to hear my opinion on certain matters. Most respondents, however, thanked me for their participation in the survey, and a few even felt the survey provided them with the necessary mechanism to vent their complaints. Many felt liberated and important that their opinions mattered. A few respondents, however, were disappointed when I explained that each respondent's observation would "only" be one among hundreds. They felt that their voices were again becoming diluted. Further, due to overall levels of fear, some people became unnecessarily worried when they were randomly selected to participate. Were they under suspicion? (Tessler and Jamal 2006, 436)

In all, she writes, she spent 3 months working 12 hours or more per day to carry out the inquiry. Given the circumstances, it seems fair to conclude that the outcome of the survey project, in terms of response rates and data quality, would have been very different without Jamal's deep engagement with respondents and with the research site more generally – if it were possible to carry out the project at all.

Working in rural China, a different kind of environment but one that also features many obstacles to political inquiry, Tsai strikes a related note in emphasizing the “socially embedded” nature of the survey research endeavor. The investigator depends on ties with local collaborators, government authorities, interviewers, and respondents alike, all of whom may be vulnerable to trouble created by the project itself. Thus, the researcher needs “to invest in building and shaping these social relationships so that they generate trust and mutual obligations.” In Tsai’s project, this meant giving interviewers time to engage in extensive conversation with officials and others in the villages that fell within the sample, in order to obtain high-quality information and cross-check discrepancies. She contrasts this with traditional approaches to survey research, which have aimed at standardizing and depersonalizing the survey process in an effort to minimize bias and error (Tsai 2010, 255, 258–260).

A researcher working in a different part of the world became dissatisfied with the results of his efforts to train research assistants to his desired levels of competence and responsibility. So he acted as his own enumerator, personally visiting sixty villages and staying half a day or overnight in each. The group discussions he held in these communities produced in-depth, qualitative data and survey data simultaneously. He explains: “I would move back and forth between asking the question on the survey and a more general discussion . . . My goal was for them to get lost in the discussion while I was collecting my data.” This approach allowed the scholar to ask key questions in a standardized fashion to each group while maintaining an open conversational style that allowed him to probe when necessary: “I would track how people were responding. If I would start to sense someone giving me artificial answers, then I would joke and be friendly and sort of point out that I knew stuff.”<sup>38</sup>

These approaches illustrate highly intensive ways of laying groundwork for large-*n* surveys. In each example, the goal was for scholars to establish their bona fides and build relationships with the population to be sampled in hopes of maximizing the likelihood of obtaining valid data. Other methods include affiliating with local authorities or institutions,<sup>39</sup> or otherwise obtaining the blessing of trusted insiders who can vouch for the researcher. Forming partnerships with local organizations can also encourage respondents to accept a survey and devote time to it. One investigator studying a grassroots

<sup>38</sup> Interview, LM-11, August 31, 2012.

<sup>39</sup> Interviews, LM-5, August 27, 2012; LM-13, September 7, 2012; LM-10, September 18, 2012.

organization distributed survey questionnaires to members of the group through a network of contacts who also were members, and obtained response rates that were “off the charts” compared to similar survey efforts in which requests to participate came directly from the investigator.<sup>40</sup> A researcher working in a conflict setting partnered with an NGO that helped to legitimate her survey. The group, which was compensated for its services, furnished office space, provided one of its staff members as a project manager, and published a newspaper article to help recruit local enumerators. The researcher and the project manager then trained and supervised the enumerators as they carried out sampling and interviews over the course of a month. This arrangement allowed the researcher to access the relevant population while retaining close operational control over the survey process.<sup>41</sup>

These examples all point to the possibility of encouraging survey respondents to share information through a sense of connection with (or an impression of legitimacy of) the researchers, or a desire to help further the project’s goals. Yet material incentives sometimes play a role as well. Political scientists vary widely in terms of how much and what type of compensation they offer to respondents, as discussed in more detail in Chapter 4. Respondents might receive cash payments, be enrolled in a lottery, be offered small gifts such as T-shirts – or may be provided no material compensation. Factors such as the length of the questionnaire, the degree of burden it imposes, and local norms shape decisions about what, if anything, should be given to respondents in return for their time and participation.

---

## Conclusion

---

This chapter has considered some of the opportunities and challenges presented by administering surveys in the field. Whether in the United States or abroad, each context presents its own distinctive possibilities and problems, many of which are not covered in textbooks and may be difficult or impossible to anticipate prior to immersion in the local setting. Political scientists, like researchers in other disciplines, have developed a wide range of strategies to capitalize on these opportunities and address these problems. They have devised methods of overcoming complications and threats to data quality posed by sampling difficulties, sensitive subject matter, inhospitable political climates, wary respondents, and more. The adaptation and perseverance that

<sup>40</sup> Interview, BR-17, November 14, 2013.

<sup>41</sup> Interview, BR-14, October 24, 2013.

doing so requires serve as a vivid manifestation of flexible discipline, a principle underlying good field research. Carrying out a survey in the field means coming up with creative ways to deal with countless details while exercising discipline in various forms – so that researchers and their teams stay on task, and so that the survey obtains valid information and will stand up to methodological critique. Doing so allows researchers to accomplish multiple analytic tasks through their survey: survey research can make multiple contributions to theory.

As noted, sometimes survey research is viewed as not fitting into a classic model of fieldwork. But as this chapter makes clear, survey work should be seen as closely related to, and often pursued in tandem with, other forms of field inquiry. Like other techniques, survey research benefits from close engagement with the local context. Indeed, such engagement is often necessary to design a survey that is feasible and meaningful in a given setting, and to encourage the forthcoming participation of those who populate the setting. Survey research is rarely the sole method of inquiry in a field research project. Rather, it is typical for political scientists pursuing the kinds of data that surveys provide to triangulate, gathering data using other techniques as well (e.g., archival research, interviews, and site-intensive work). Indeed, quantitative data gathering can be integrated with, can be informed by, and can inform a project's qualitative dimensions in many ways. This iterative knowledge-generation, we believe, is one of the hallmarks of field research, and one of the qualities that make it such a powerful form of empirical inquiry.

Yet the point is not just that surveys and other modes of research enrich one another. We also wish to emphasize that, on the ground, in action, the seemingly intuitive distinction between surveys (impersonal, quantitative, involving brief interactions that produce "thin" data) and other field techniques (high-touch, qualitative, involving longer periods of contact that produce "thick" data) often breaks down, partially or entirely. Survey work often requires forms of immersion in field sites. Encouraging respondents to give candid, illuminating answers to certain kinds of questions may require building understanding, identification, and trust in various forms. The information produced through surveys may, when appropriate, be rich and detailed, even if the goal is later to code that information in structured, quantitative forms.

More than any of the other data-collection techniques considered so far (but like field experiments, which are addressed in the next chapter), survey work often necessitates collaboration. This brings out new aspects of the PI's

“project manager” role, as discussed in Chapter 4. As the examples above have shown, whether the PI/PM is working with a single assistant, training whole teams of enumerators, or contracting out some of the work to a survey organization, she generally needs to remain as present, engaged, and vigilant as possible. Bringing other people and groups into the project generally means not just training but also motivating collaborators to approach the survey process with appropriate care and with commitment to treating respondents ethically and respectfully.

This chapter has emphasized the ways in which researchers adapt their work to the particularities of a field site. Yet by illustrating such adaptation, we in no way wish to suggest that rigorously preparing for survey research is not critical. Surveys are in no sense a form of research in which “flexibility” should overshadow “discipline” – quite the contrary. What scholars are adapting, we insist, should be a project designed according to general principles of survey research. Projects designed with those principles in mind have firm foundations – and will produce results that make sense to, and are persuasive for, scholarly audiences. Even as a researcher grapples with the challenges of the field, sometimes what she needs most comes from back home. As one experienced investigator said, reflecting on his dissertation research: “The main thing I wish I had going into it was I wish I had done more homework in the academic literature side of things. For example, I wish I’d read a lot more about panel studies, the nuts and bolts about questionnaire design, question wording. I wish I’d been more familiar with debates on partisanship and key concepts. I came back later, and said ‘I have decent data but I wish I’d measured it *that* way.’ . . . People need to read, read, read stuff that’s there so they don’t have that moment when they kick themselves.”<sup>42</sup>

<sup>42</sup> Interview, BR-3, August 6, 2012.

## Experiments in the field

Even as recently as the late 1990s, experiments occupied a small and relatively marginalized niche in political science research. Despite early precursors such as Harold Gosnell's (1926) study of voter turnout in the 1920s and Samuel J. Eldersveld's (1956) study of propaganda and voting behavior in the 1950s, experimental research was traditionally considered to be impractical, if not impossible.<sup>1</sup>

Today, by contrast, experimental methods are both more prevalent and more widely respected than in the past. Druckman *et al.* (2006) found that more articles based on experimental data are being published in top-tier journals, and that they are being cited significantly more frequently than comparison groups of articles. Political scientists have recently published a number of noteworthy books and journal articles on the methodology of experimental research, as discussed below. In addition, more and more graduate programs have added seminars on experimental methodology, and intensive summer training courses have incorporated sessions on the topic.<sup>2</sup> Finally, the creation in 2010 of an APSA Organized Section on Experimental Research, and its introduction of the *Journal of Experimental Political Science* in 2013, provide further evidence that experimental

<sup>1</sup> Lawrence Lowell, in his APSA presidential address (1910), advised scholars in the nascent discipline of political science to eschew the model of the natural sciences contending that: "We are limited by the impossibility of experiment. Politics is an observational, not an experimental science." Over sixty years later, Lijphart (1971) described the experimental method as "nearly ideal" but rare due to the "practical and ethical impediments." One overview of the history of experiments in political science emphasizes the ongoing efforts of generations of researchers, holding that this history is "more substantial than most recognize" (Morton and Williams 2010, 3–27, at 9).

<sup>2</sup> For example, the ICPSR includes a week-long session on designing and conducting field experiments. The Empirical Implications of Theoretical Models (EITM) workshop at Washington University has included experimental methods. And the IQMR has sessions dedicated to natural and field experiments. Experiments in Governance and Politics Network (EGAP), a group focusing on experimental work in the political economy of development, holds meetings and maintains a web site (<http://egap.org>).

methods, and empirical and methodological research based on them, are becoming broadly institutionalized in the discipline.<sup>3</sup>

Findings from the FRPS survey reflect both the relative paucity of field experiments overall since the 1960s, and their recent upswing in usage.<sup>4</sup> Experiments were employed in only 2.8 percent of all reported projects begun from 1955 through 1999. By comparison, focus groups, another relatively uncommon data-collection technique, had more than three times this prevalence in the same time period; surveys were more than eleven times as popular. But experiments have surged. Among projects that our survey respondents began between 2008 and 2011 (the last year covered by the survey), 10.5 percent included experiments. While still only one-third as prevalent as ethnography and participant observation, for instance, experimental research is on the rise. Experiments are about equally common in projects within the United States and those involving international locations, and are used by political scientists of various subfields, including international relations and public administration.

Several developments have likely driven the recent increase in prominence of experiments in political science. The experimental method's ability to provide strong evidence of, and accurately estimate, causal effects, as discussed below, is central to its appeal. Along lines discussed in Chapter 2, influences from other disciplines have also played an important role in encouraging the diffusion of the technique. Laboratory experiments studying political behavior and cognition have been broadly influenced by work in social psychology and behavioral economics. Field experiments by political scientists, meanwhile, received a shot in the arm from the work of developmental economists and institutions associated with them, such as Innovations for Poverty Action (founded 2002) and the Abdul Lateef Jameel Poverty Action Lab at MIT (founded 2003).<sup>5</sup> Indeed, discussions of economists' experimental work have even appeared in magazines and books intended for mass audiences outside of academia.<sup>6</sup> Within political science, experiments have been championed by vocal figures at highly respected

<sup>3</sup> The APSA section on Experimental Research counted 409 members in September 2012. As a reference point, the largest APSA section included 1,351 members (Comparative Politics) and the smallest had 183 members (Canadian Politics) in the same time period.

<sup>4</sup> These figures capture projects in which either the use of "field experiments" or the use of "laboratory experiments" was reported, with the latter indicating what we refer to below as "laboratory experiments in the field."

<sup>5</sup> [www.poverty-action.org/about/story](http://www.poverty-action.org/about/story), accessed December 5, 2013.

<sup>6</sup> See the Parker feature on Esther Duflo and the Poverty Lab (2010) as well as Banerjee and Duflo's book (2011).

institutions. The application of experimentally tested findings in practical politics, for instance in Barack Obama's successful presidential campaigns and through organizations such as The Analyst Institute, have likewise raised their profile.<sup>7</sup> Finally, the growth in experimental research is also driven by a demand for rigorous evaluation of policy interventions by donors, foundations, governments, and non-governmental organizations (Humphreys and Weinstein 2009). All of these theoretical, methodological, and policy innovations have been facilitated by advances in technology, such as the development of computer-assisted telephone interviewing and related software programs (Druckman *et al.* 2006; Sniderman 2011).<sup>8</sup>

As with other research traditions and data-collection techniques considered in this book, the social science literature on field experiments, and on experiments more generally, contains a number of overviews of this technique. Some focus on the technical dimensions of experiments and on analyzing experimental data. We do not here address the fundamentals of experimental design and analysis of experimental data, though we refer readers interested in these topics to appropriate sources.<sup>9</sup>

Instead, and as in the other chapters of this book that focus on data-collection techniques, our main goal is to *contextualize* experimental research in the field. We examine the effect field circumstances have on the deployment of experiments, explore the challenges of using experimental techniques in the field, and offer strategies for addressing them. Throughout, we demonstrate that engagement with context (a core principle of good field research) is at least as important in this research tradition as in others: field experiments are – and must be – grounded in detailed knowledge of the setting and the subject population. Indeed, field experimenters often engage in considerable periods of open-ended study as they conceive and

<sup>7</sup> See, for instance, Benedict Carey, "Academic 'dream team' helped Obama's effort," *New York Times*, November 12, 2012, and Issenberg (2012). The Analyst Institute is "a clearinghouse for evidence-based best practices in progressive voter contact," which "assists organizations in building testing into their voter contact efforts." From <https://analystinstitute.org/>, accessed December 5, 2013.

<sup>8</sup> In addition to the published literature, this paragraph draws on two interviews: BR-17, November 14, 2013 and BR-18, November 18, 2013.

<sup>9</sup> On the specific topic of field experiments, Gerber and Green's text (2012) provides a particularly valuable resource, explaining in detail the process of estimating causal effects while coping with technical complexities in a variety of scenarios. Other important book-length resources on experiments in general are Morton and Williams (2010), Mutz (2011), and Druckman *et al.* (2011a). Gerring (2012, Chapters 9 and 10) provides one explanation of the logic of randomized designs in causal inference. Books with a weaker connection to political science per se that address experimental methods include Cox (1958), Shadish, Cook, and Campbell (2001), and Maxwell and Delaney (2004).

plan their work, and the core ideas behind good experiments often emerge from extended contact with particular locales.

We begin by identifying the key rationales for experimental research, discussing how it differs from observational research, and describing various types of experimental methods. Next, we examine the analytic tasks that can be accomplished by carrying out experiments in the field. In contrast to the chapters focused on observational data-gathering techniques (Chapters 5 through 8), we do not offer an extended discussion of how experiments contribute to theory building by helping scholars achieve multiple analytic goals (e.g., formulating research questions, selecting cases, conceptualizing and measuring key variables, generating potential explanations, and illuminating causal processes and mechanisms). The reason is that experimental methods are generally understood to have a specialized strength and relatively specific application: determining the effect on a given outcome of a particular causal treatment (i.e., determining the causal relationship between a particular “X” and a particular “Y,” or identifying “effects of causes”).<sup>10</sup> Put in terms of the analytic tasks we discuss in other chapters, we might say experiments are generally and primarily aimed at hypothesis testing. Further, and as we discuss in more detail below, experiments are often conducted at a relatively advanced stage in a given research project. Given this timing – and their elaborate nature and the potentially high cost of carrying them out – scholars generally have laid much of the analytic groundwork for their projects when they reach the point of executing an experiment. We limit our discussion of experiments’ analytic contributions, then, to addressing their ability to illuminate specific causal relationships.

As one might expect, given this analytic specialization, experiments are typically combined with the use of other data-collection techniques and analytic methods, and often form one facet of broader research programs that extend through multiple studies. The way in which political scientists draw on interviews, surveys, and ethnography or participant observation – prior to, during, or after carrying out an experiment in the context of any particular study – are the topics of the chapter’s subsequent section. Combining multiple methods helps scholars to capitalize on the field context to accomplish a broader range of analytic tasks, and helps to mitigate the risk

<sup>10</sup> Correspondingly, experiments are less apt for identifying the “causes of effects.” As one experimentalist put it, “If you’re trying to explain all the variance in a behavior, an experiment is not suitable for that. For example, why do people turn out to vote? There are so many different reasons. That’s not best addressed by an experiment. Maybe just a survey.” Interview, BR-15, October 25, 2013. Gerring discusses the distinction between effects-of-causes and causes-of-effects analyses (2012, 333–335).

that the experiment will go awry or simply produce uninteresting results. In short, triangulation with other methods – another core principle of good field research – figures prominently in the process of field experimental research.

The final four substantive sections of the chapter address the practical processes of conceiving, developing, and carrying out experiments in the field. While reading published articles based on experimental studies allows one to reverse-engineer, to an extent, the steps the researcher took, such work almost inevitably omits much information about the context and backdrop for the experiment, and many practical details. A handful of reflective and prescriptive essays by practitioners make these steps more explicit and distill general lessons for others to follow.<sup>11</sup> By drawing on this material, and on interviews with other practitioners about their experiences, we provide practical suggestions on how to undertake this form of work efficiently and appropriately. We first consider the benefits of partnering with local organizations such as NGOs or research institutions in the conduct of a field experiment, emphasizing how developing and managing such collaborations require knowledge of the research setting and the strengths and weakness of specific organizations. We then discuss the development, randomization, and application of treatments, as well as dealing with subjects.

Precisely because the experimenter applies a treatment to certain units or subjects, seeking to evaluate the change that is brought about as a result, this form of research raises ethical questions that are absent or less prominent in other modes of research. While most forms of inquiry affect human subjects in some way, field experiments contain one or more elements that call for special ethical scrutiny. These include the deliberate introduction of change in the form of a treatment, the integration of real-world political events and phenomena into the experiment,<sup>12</sup> and (often) the application of treatments on a large scale. Issues are thus raised about obtaining informed consent from those involved in the experiment; the possibility of influencing or

<sup>11</sup> Sources of this kind are cited throughout the chapter, and we offer just a few examples here: Gueron (2002), Duflo, Glennerster, and Kremer (2006), Humphreys and Weinstein (2009), Paluck (2009), Humphreys (2011), as well as some of the extended examples in Gerber and Green (2012).

<sup>12</sup> For example, a field experiment conducted by Leonard Wantchekon (2003) was embedded in the context of the 2001 presidential election in Benin. The experiment, conducted in collaboration with campaign managers from major parties, involved exposing residents of certain villages to political appeals based on clientelism, and others to appeals based on policy programs. Wantchekon chose only villages in “safe districts” where certain candidates had such a stronghold that they were very likely to win and thus argued that the experiment could not have changed the result of the election (De La O and Wantchekon 2011).

disturbing important political processes; deceiving participants; and exposing participants to tangible or intangible harms or depriving them of benefits. These matters – all relating to the principle of ethical commitment – are addressed throughout the chapter.

## Experimental methods: rationales and types

The theoretical rationale for experimental methods in the social sciences grows out of the goal of understanding causal relationships. In observational studies – those that involve observing or measuring social phenomena as they are found – causal inference can be flawed or questionable, and the direction of causality left uncertain. It can also be difficult to rule out the possibility that unmeasured confounding factors, including selection effects, might bias the estimates of coefficients on explanatory variables.

Experiments, by contrast, hold the promise of offering rigorous tests of specific causal effects. Their advocates note that the challenges to causal inference that haunt observational studies are essentially removed through the design of experimental research, which proceeds as follows. Hypotheses are developed *ex ante*. Variation in the hypothesized causal factor(s) is not left up to “nature,” but rather is determined by the researchers themselves, and instantiated in the experiment in the form of different levels or types of treatment. Researchers assign subjects to treatment group(s) and a control group *at random*. A specific manipulation or treatment is applied to the former, and the groups are later compared with respect to outcomes of interest. The randomization makes it possible to assume that the two or more groups are effectively identical in all theoretically relevant aspects except for the applied treatment (and this assumption can be checked by comparing observable measures across the groups). Thus, any statistically significant differences in outcomes are presumed to be attributable to the treatment. Given the above, some consider experiments to be the “gold standard” of empirical work, although others take exception to this characterization.<sup>13</sup>

Experimental field research differs from other types of field study in fundamental ways. Most forms of research considered in this book aim to understand the political world through observational means. Causal

<sup>13</sup> See the debate between outspoken advocates such as Duflo *et al.* (2006) and critics such as Deaton (2009). This controversial debate and the notion of a “gold standard” were referred to in the interviews we conducted in connection with this book as well.

processes are studied – perhaps as they are playing out, perhaps after the fact – but they are not created by the investigator herself. Researchers who conduct interviews or surveys do not speak of applying a treatment to those whom they study. Indeed, researchers often strive *not* to affect, manipulate, or disturb their research subjects (as in much site-intensive work) or else are unable to do so (as in archival studies). While interviewers commonly attempt to stimulate or provoke a response through the questions they ask, their goal is generally not to assess the causal effect of the questions themselves. Rather, interviewers hope that the way in which study participants answer the questions they pose will help them to understand and identify the causes of the phenomenon of interest.

The nature of assignment to experimental conditions also differentiates true experiments from what are called natural experiments, or quasi-experiments. In natural experiments, the researcher does not assign the treatment himself, but rather relies on pre-existing processes that have created random or “as-if random” assignment. An example of a natural experiment is Chattopadhyay and Duflo’s (2004) study of the effects of the Indian government’s quota system for women’s political participation on the provision of public goods. The authors exploited the introduction of a 1992 constitutional amendment stipulating that one-third of village council chief positions be reserved for women; in at least two states, villages were chosen to have such a reservation through random assignment. They found that expanding female political representation made access to public goods more equal. In research based on natural experiments, then, the researcher comes upon random or near-random processes, generally after the fact, rather than creating them herself. We thus limit our treatment of natural experiments to a discussion of how field research can help scholars to discover natural experiments.

Gerber and Green suggest that experiments can be thought of as varying in their degree of “fieldness” along four dimensions: contexts, participants, treatments, and outcome measures (2012, 11). On each of these dimensions, a given experiment could be more artificial and removed from populations of interest, or closer to the actual experiences of people in the real world. At one end of the spectrum, a purely non-field experiment exploring voters’ responses to political advertisements might take place in a university computer lab, with college students viewing one of two statements from a hypothetical candidate on a computer screen and subsequently registering a response to a question or prompt via a mouse click. In fact, a prominent strain of experimental work in political science (one often used in political

psychology)<sup>14</sup> takes place in laboratory settings. In such studies, the researchers have maximal if not complete control over all aspects of the experimental context. They determine whom to recruit as subjects; subjects participate in the experiment in a location chosen by the researchers; the investigators dictate what the subjects see, hear, and experience, including the experimental treatment; they control or prevent interactions with other participants; and they can generally ensure that relatively few subjects will leave in the middle of the experiment rather than completing it. In the middle of the spectrum might be laboratory-in-the-field experiments, discussed below. And at the other end of the spectrum, a purely field-based study might be set in an actual city council election, with real voters receiving in their home mailboxes one of two versions of a campaign message from a real candidate; here the outcome of interest might be subjects' actual vote choice on election day.

Field experiments can be carried out in myriad types of contexts, leading us to consider the multiple roles and meanings of the field. Indeed, conducting an experiment in the field means incorporating into the study some of the complexity, particularity, and richness that characterize actual political settings. To extend the election example, the field study would entail engaging with the politics of an actual city, perhaps buffeted by a budget crisis or a police-brutality scandal in the course of the campaign; it would mean integrating the dynamics of an election cycle, over the course of which citizens' perceptions fluctuate in response to debates, media coverage, and other factors. Field experiments often mean road-testing a particular intervention that is intended by policymakers to be beneficial in some way, as in a program to give voters more information about candidates or to bring women into leadership positions that have historically been closed to them. Such interventions are subjected to real-world stresses and challenges, and their effectiveness rigorously evaluated. To the extent that experimentalists opt for real-world settings, their work suggests that it is not sufficient to study the topics of interest in clinical, laboratory conditions and in ways that are removed from the actual setting of politics.

Field experimentalists inherently aim for a degree of naturalism, in part as an effort to increase the external validity of their study: the ability to generalize the results.<sup>15</sup> In some experiments, participants may not even

<sup>14</sup> For instance, see Redlawsk, Civettini, and Emmerson (2010).

<sup>15</sup> Gerber notes that field experiments "aim to reproduce the environment in which the phenomenon of interest naturally occurs and thereby enhance the external validity of the experiment" (2011, 116).

be aware that they are taking part in a study, and thus they may behave as if unobserved, eliminating forms of bias that stem from subjects' awareness that they are under scrutiny. Rarely is this possible when scholars deploy the other data-collection techniques involving human subjects that are discussed in this book, unless researchers engage in some degree of deception. For instance, interviewees are generally fully cognisant that their answers to the questions they are asked are being evaluated and analyzed by their interlocutor. Even in participant observation – a technique scholars may adopt, at least in part, in hopes of blending in to the field context to some extent – researchers' presence can subtly but consistently remind the people with whom they are interacting that they are being studied.

But naturalism also has potential costs, and can increase the number of challenges experimentalists face in conducting their study. Utilizing real-world settings rather than a laboratory, studying actual populations of interest rather than proxies, and extending the study beyond a single sitting or interaction tends to mean sacrificing a degree of control over the experimental milieu and accepting increased complexity and risks to the study's integrity.<sup>16</sup> Humphreys and Weinstein note that, in field as opposed to laboratory experiments, "Features such as the characteristics of subjects, the information available to them, and the precise manner and context in which the treatment is applied are more likely to take on values given by 'nature' rather than being set at the discretion of the investigator" (2009, 369). Participants who are supposed to remain independent may interact with one another instead. Units that are assigned to receive a treatment may end up not receiving it. And subjects may drop out in the middle of an experiment.

As suggested above, a "laboratory-in-the-field" study is a kind of compromise between the pure-laboratory and pure-field extremes. Such studies entail laboratory-like arrangements that are brought to a researcher's field site(s) in order to examine particular populations or naturally occurring situations (Morton and Williams 2010, 296). Such a study involves enhanced verisimilitude on one or more of Gerber and Green's (2012) four "dimensions of fieldness" (contexts, participants, treatments, and outcome measures), but not all of them. For example, some researchers might conduct similar experiments in university classrooms in two or three different

<sup>16</sup> For Morton and Williams (2010, 46), in field experiments, "a researcher's intervention takes place in subjects' natural environments and the researcher has only limited control beyond the intervention conducted."

countries, recruiting participants locally in each setting. Even if no other aspect of the experiment has a “field” quality to it, such a project would be substantially different from one conducted in a laboratory at a scholar’s home institution, perhaps relying for subjects on the typical American “college sophomore.”<sup>17</sup> An example of such a laboratory-in-the-field experiment is found in the work on cooperation between co-ethnics in laboratories in Kampala, Uganda (Habyarimana *et al.* 2007, 2009). The objective of these studies was to disentangle various potential causal mechanisms for co-ethnic cooperation. Other lab-in-the-field studies test theories developed in laboratory environments in the United States in additional contexts. This was the primary objective for experimental work conducted on the dynamics of individual decision-making regarding common-pool resources with local populations in forty-one countries (Ahn, Ostrom, and Walker 2010).

In short, experiments come in various forms, each with characteristic strengths and weaknesses, different levels of control, and diverse sorts of interaction with “the field.” This variation notwithstanding, much experimental research has the same goal: rigorously assessing a single causal relationship or a small number of causal effects. The chapter’s next section examines this singular analytic focus.

## Field experiments’ analytic contributions

Previous chapters addressing other data-collection techniques such as interviewing or site-intensive methods have highlighted their versatility, demonstrating how they can be employed throughout the research cycle to aid in accomplishing multiple analytic tasks, including sharpening questions, refining concepts and measures, generating hypotheses, and adjudicating among competing hypotheses. Experiments, by contrast, are typically seen as having a more specific application: assessing a particular cause-and-effect linkage. As such, they are generally carried out after scholars have already refined their questions, and potential explanatory hypotheses. Of the various data-collection techniques we consider, field experiments may be the most narrowly focused, yielding a very specific type of information about politics in

<sup>17</sup> While many experimentalists use student subjects in their work (Kam, Wilking, and Zechmeister 2007), scholars debate whether this is a significant problem for external validity. While Sears (1986) raised doubts about the validity of conclusions based on student participants, Druckman and Kam (2011) argue that student subjects do not “intrinsically” undermine a study’s external validity.

the field context. For experimentalists, this is balanced by the strength of the evidence that experiments provide for understanding causal relationships.

The causal linkage of interest in a particular experiment may be considered important for several reasons. It may play an important role in a body of theory. It may have practical or policy-related implications, as when researchers determine whether distributing mosquito netting prevents the transmission of malaria, or whether electronic voting machines reduce spoiled ballots. Often the issues at hand are both practical and theoretical, as in research on citizens' decisions whether or not to vote. As Roth (1995) puts it, experiments have three broad types of objectives: "searching for facts," "speaking to theorists," and "whispering in the ears of princes" (that is, aiming to identify policy implications).

Field experiments in political science vary tremendously with regard to topics studied and treatments applied. Voting behavior and voter turnout have been prominent areas for field experiments. In such studies, voters are typically given a treatment in their homes by being contacted by canvassers, phone callers, or mailings, producing findings that have emphasized the role of contact and social pressure in political behavior (Gerber and Green 2000; Michelson 2003; Gerber, Green, and Larimer 2008; Nickerson 2008; García Bedolla and Michelson 2012). Yet experiments examine many other topics, and significant innovation has occurred with regard to approaches to treatments. In research on post-genocide Rwanda, for example, Paluck and Green (2009) studied the question of whether it is possible to change cultural norms of deference to authority. They found that residents of seven communities who listened to a radio program designed to promote independent thinking tended to internalize that message and were more willing to express dissent than were members of a control group.

Advocates maintain that many phenomena that have been studied observationally are amenable to randomized controlled trials. Many field experiments deal with things that individuals do – in other words, micro-level political behavior. Yet with the cooperation of governments, many realms of public policy and larger-scale dynamics and trends – economic growth in small towns, or corruption among government contractors, for example – can potentially be examined through randomized studies. Even so, there are limits to the kinds of questions that can be addressed using experimental techniques. Experiments require a treatment that can be assigned. However, fundamental aspects of individual identity, such as (perhaps) a person's being a New Yorker as opposed to a Minnesotan, or of political structures, such as a country's having a parliamentary or a presidential system of

government, can be controlled for but cannot realistically be randomized.<sup>18</sup> There is also the requirement that units in the treatment and control groups remain independent from one another during the study – a criterion that can be difficult to sustain in many kinds of interconnected, communicative populations. Apart from these basic realities, many ethical as well as practical considerations, as discussed below, constrain the options available to scholars who contemplate an experiment.

Given the excitement surrounding field experiments and their rise in popularity, graduate students often clamor to do one. Experimentalists highlight the importance of carefully evaluating the challenges involved before launching such a project, however. For example, in an interview, one scholar cautioned graduate students and junior faculty that field experiments are “inherently very risky” because “a lot of them fall apart.”<sup>19</sup> Much of the advice in Chapter 8 concerning the time costs and logistical challenges of surveys also applies to experiments. One interviewee was surprised by the extensive managerial work entailed and concluded that it was “not my strong suit.”<sup>20</sup> Thinking through the details of the multiple steps involved in conceptualizing, designing, preparing, and carrying out an experiment, and considering the successes and failures that other researchers have encountered in the field, helps to build a realistic understanding of what it takes to employ this technique effectively.

In sum, experiments are versatile, but subject to limitations, with regard to the topics they can be used to study. Their analytic contributions tend to be comparatively focused. Accordingly, the design of field experiments requires a great deal of analytic precision, and thus they are often most fruitfully employed after critical research design choices have been made. Moreover, carrying out field experiments can be costly, challenging, and risky. For these reasons and several others outlined in the next section, experiments are often combined with other analytic techniques and methods in order to accomplish the objectives of a given research project or agenda.

<sup>18</sup> As Humphreys and Weinstein put it: “Many political processes and attributes are likely to remain (and likely ought to remain) unavailable for experimental manipulation: whether governments are authoritarian or democratic, whether regions secede, whether governments launch brutal counterinsurgency campaigns, whether a given individual adopts a given set of preferences, etc.” (2009, 374).

<sup>19</sup> Interview, LM-13, September 7, 2012.

<sup>20</sup> “One of the things I didn’t quite realize before grad school is that people really vary in their talents that are required for field research. I discovered I’m not that great at the management side. I was more on the data side.” Interview, BR-16, October 31, 2013.

## Combining experiments with other methods

Field experiments often emerge out of, and form part of, larger research programs involving the use of other data-collection techniques and analytic methods, both before and after an experiment has been conducted. Indeed, the fact that political scientists often employ other kinds of methods in concert with experiments is clearly reflected in the FRPS survey data. Of the 66 field research projects reported in the FRPS survey that employed experiments, only 3 involved no other data-collection techniques. Of projects including experiments, 73 percent also involved the use of surveys; 77 percent involved interviews; and 48 percent involved ethnography or participant observation. Indeed, projects employing experiments also used, on average, nearly 6 non-experimental data-gathering techniques. Almost three-quarters included either qualitative or interpretive approaches to analysis.

In short, while the survey data do not allow us to distinguish the particular role played by experiments versus those played by other methods in the course of a given project, they do make clear that the combination of multiple methods is common among experimentalists. The first part of this section discusses several reasons why this is the case – although the benefits of thorough and multi-method engagement with one's field site form a theme that recurs throughout the chapter. The second explores *how* experiments can be combined with other forms of inquiry.

### Why combine methods?

The previous section offered several reasons why field experiments are often carried out in tandem with the use of other research techniques – often *after* those techniques have been used to generate significant knowledge of the experimental context. To reprise briefly, experiments are generally best employed to nail down particular causal relationships, meaning that alternative data-collection techniques and analytic strategies must be used to achieve other core research objectives. Moreover, because experiments can be expensive in terms of time and money, researchers need to be sure they are designed and carried out as effectively as possible. Also, experiments require a high level of analytic clarity. Often, this confidence and clarity only develop later in a project when a scholar has learned a great deal about the field site through analyzing data collected via other methods. One of our interviewees offered the following thought on the issue of sequencing experimental and observational data collection:

I don't think it necessarily has to be that the experiment comes last. But it's hard to imagine the experiment coming out of nowhere. Experiments are usually expensive and time-consuming to run, so people only want to do ones that have a chance of working. That can only happen after either a process of gathering hypotheses that have been established in other literature, or a lot of fieldwork of your own.<sup>21</sup>

Put somewhat differently – like all of the data-collection techniques addressed in this book, experiments (even laboratory studies that are intended to be as generic and generalizable as possible) are employed at a certain time in a certain location, and involve a certain kind of subject pool with particular experiences, cultural standards, norms, ideas, and assumptions. They investigate the effect of a certain kind of mailer or phone call on voter turnout in a certain election in a certain year, for instance, or the influence of a certain kind of technology on women's participation in villages of a certain region of a certain country. The more natural the setting, the more important it is that researchers have a deep understanding of the context in order to design and conduct an appropriate experimental intervention – an understanding of the type that often only comes through previous research on that setting. As Paluck (2009) notes, many essential decisions relating to the conduct of field experiments – some with the potential for life-or-death consequences in zones of high insecurity – emerge from an intimate knowledge of the field context.

In fact, the very idea for a field experiment often emerges out of immersion in a field setting. Identifying worthy topics that are amenable to experimental study – and finding and developing fruitful working relationships with partners – may require long-term engagement with a locale and the questions or problems that it presents. And on a practical level, conducting an experiment – arranging transfers of funds, obtaining approvals, bringing in necessary equipment, timing the roll-out of the study – often requires considerable familiarity with a field site. One example can be found in Olken's experimental studies in Indonesia, examining such topics as the effect of audits and participatory grassroots monitoring on corruption in road-construction projects in 608 villages, as well as the effect of direct democracy and representative democracy on development projects (2007, 2010). Carrying out these studies required deep cooperation with the Indonesian government and one of its development programs. It also required developing a thorough grasp of Indonesia's political history, village conditions in multiple provinces of the country, and the practical workings of both village governance and the

<sup>21</sup> Interview, BR-17, November 14, 2013.

financing of construction projects.<sup>22</sup> Paluck (2009) also acknowledges her great appreciation for a primary research team leader from the local area who provided vital feedback throughout the research process.

In addition to helping a researcher learn enough about a context to conduct an experiment there effectively, integrating experiments with other research methods helps to ensure that a research project will pay off. As previously noted, field experiments can fall apart for various reasons, or simply find no support for the hypothesis in question.<sup>23</sup> Several interviewees made the point that pursuing a topic through multiple methods helps to hedge against such occurrences. As one said: “A field experiment should be one component to a research project, but it’s most effective when it’s embedded in a larger kind of research project with lots of components: qualitative, quantitative, descriptive. So that if it [the experiment] works out logically and it maps onto interesting theories, that’s fantastic, but it doesn’t all depend on that.”<sup>24</sup>

Finally, it is often pointed out that experiments allow for particularly high internal validity (confidence that the treatment truly has the causal effect it is claimed to have) yet leave external validity (the generalizability of findings) an unanswered question. For experimentalists, a primary solution to the external validity problem is to conduct variations of the same experiment, in the same locale or similar ones, in order to build deeper knowledge of causal relationships. One example of this approach is provided by successive studies of get-out-the-vote (GOTV) techniques in the United States.<sup>25</sup> Researchers can also replicate experiments in other settings in order to test the generalizability of a causal effect, for instance in research on the effects of mandated participation by women in local governance.<sup>26</sup> Yet knowing where and how to carry out follow-on studies that can best advance the research program requires understanding the practically and theoretically relevant characteristics of a given field context – so future experiments can be

<sup>22</sup> Olken was affiliated with the World Bank’s Jakarta office as a consultant for seven years, and speaks Indonesian. See the cited articles as well as a profile of Olken, “Graft paper,” accessed November 17, 2013 at [www.american.com/archive/2008/january-february-magazine-contents/grafft-paper](http://www.american.com/archive/2008/january-february-magazine-contents/grafft-paper).

<sup>23</sup> It is often pointed out that negative results, too, can be important contributions to knowledge. Yet, in what is known as “publication bias,” journals are less likely to publish such work.

<sup>24</sup> Interview, BR-16, October 31, 2013.

<sup>25</sup> For instance, García Bedolla and Michelson (2012) report findings from experiments that build and expand on earlier GOTV research.

<sup>26</sup> One example is research by Ocantes *et al.* (2013), replicating the same survey experiment about vote-buying in several Latin American countries. Another is Beath, Christia, and Enikolopov’s study of village development councils in Afghanistan that mandate women’s participation, building on related work elsewhere (2013).

designed to vary and test these characteristics.<sup>27</sup> In other words, enhancing external validity in experimental research entails significant knowledge of multiple contexts – knowledge that often can only be generated through research in the prospective field sites.

To be clear, this is not to suggest that all field experiments require long-term immersion in the field. As we have sought to suggest, the key fieldwork principle of “engagement with context” is not perfectly correlated with duration of stay. Sometimes an experiment requires relatively little time in the field, though this may be particularly common when the location of the experiment is close to the investigator’s home institution and thus already familiar (for example: Han 2009b). Our point is simply that carrying out additional forms of inquiry in tandem with conducting field experiments allows researchers to gather information about the field context that is critical to setting up an effective experiment – whether one is carrying out the initial test of a hypothesis, or seeking to replicate a previous experimental finding. Gathering data using other techniques and employing other methods in combination with experiments also facilitates triangulation and verification – and serves as an insurance policy, ensuring that one’s research will be valuable even if the experiment should be derailed.

### **Ways and means of combining experiments and other techniques**

Many experienced field experimentalists attest that learning extensively about subjects through observational methods is often an essential prerequisite for developing meaningful experiments – yet observational methods can also be fruitfully employed in parallel with, or as follow-ups to, experiments. This subsection discusses several ways in which observational and experimental techniques can be combined and sequenced.

Intensive field-based study using observational techniques can lay the groundwork for experimental research. A process of inquiry involving interviews or site-intensive methods, for example, might identify what seems to be a causal relationship that a researcher can then assess through a field experiment. One scholar working in the developing world stressed the importance of inductively generating hypotheses by watching the work of NGOs and having open-ended conversations with members of the communities under study: “If you’re going to study this population, you can’t generate ideas for effective treatments unless you’re actually doing the

<sup>27</sup> Interview, BR-16, October 31, 2013.

soaking and poking. It is a high-labor-for-reward situation.”<sup>28</sup> According to this researcher, hypotheses derived purely from theoretical expectations in the abstract often fail to find empirical support when they are “parachuted in” for experimental testing. Economists echo such themes. One field experimenter working on agriculture in Kenya noted that interviews and focus groups enabled his team to determine plausible hypotheses explaining their puzzle: why maize farmers used little fertilizer despite the large returns it produces.<sup>29</sup> On the basis of what they learned through this qualitative work, they proceeded to test interventions aimed at lowering barriers to investment in fertilizer.

Yet the ways in which other, generally qualitative, field research techniques augment and facilitate experimental projects go well beyond merely providing a causal hunch. Immersion and qualitative methods also contribute to the core of the experimental design. Dunning explains that, in his co-authored study of fictive kinship or “cousinage” in Mali, extensive interviewing was necessary in order to determine the relationships among multiple ethnic and cousinage groups. These formed a matrix that was used in the random assignment of subjects to treatment categories, a matrix that was tested through initial trials of the experiment and refined in iterative cycles of further interviewing.<sup>30</sup>

Moreover, non-experimental techniques can be sequenced during and after an experiment has been conducted to help researchers to interpret its results. While experiments can provide powerful evidence for or against the existence of the hypothesized causal relationship, other questions often remain to be answered. As one experimentalist put it: “The big natural question is *why*. Field experiments are terrible at telling you why, they just tell you whether or not.”<sup>31</sup> That is to say, merely finding a statistically significant average treatment effect does not, in and of itself, explain what brings about that effect. And if no effect is found, there could be many

<sup>28</sup> Interview, BR-2, July 30, 2012.

<sup>29</sup> Jonathan Robinson, talk at Center for Global International and Regional Studies, University of California, Santa Cruz, March 5, 2012.

<sup>30</sup> As Dunning writes: “The point is that eliciting a reliable map of cousinage relations from key informants very centrally involved qualitative as well as mixed methods. For instance, to revise our cousinage matrix we conducted qualitative interviews with key informants. We then also employed quantitative analysis of the experimental data from initial trials. To improve the cousinage matrix, we therefore iterated between focused interviews, new versions of the cousinage matrix, and our experimental data to improve the random assignment mechanism in this experiment” (2008, 22). Findings from the study are reported in Dunning and Harrison (2010).

<sup>31</sup> Interview, BR-16, October 31, 2013.

reasons why that is the case – and understanding those reasons is likely to be of great importance for designing a follow-up experiment.<sup>32</sup> While it is possible to build into the experimental design means of addressing “why” or “why not” questions, qualitative and non-experimental quantitative methods can also augment randomized controlled trials, as experimentalists have eloquently described.

For instance, surveys or interviews may be employed in parallel or in sequence to complement the experiment itself, to shed light on the causal relationships in question and the causal mechanisms at work, to interpret its results, and to discover further hypotheses to test (Sherman and Strang 2004; Paluck 2010). Sherman and Strang (2004) advocate the use of ethnography during the course of experimental work, in part as a way of understanding the “why” behind a causal effect. In their example, a randomized controlled trial evaluates a “restorative justice” program in which crime victims meet with the offenders who harmed them in a conference led by a facilitator. The experiment determines whether, for example, victims who participate in such conferences are less likely to experience post-traumatic stress disorder and more likely to hold normal jobs. Ethnographic work proceeds in parallel, with researchers repeatedly visiting program participants over time to gain a deep understanding of their life circumstances and their emotional and psychological responses to the program. The authors maintain that this is useful not only for explaining the mechanisms at work, but also for discovering important but not self-evident hypotheses to test.<sup>33</sup>

While this chapter focuses on true experiments, involving randomized assignment by the investigator to treatment conditions, it is worth noting that field research can also contribute to the discovery of natural experiments. Obtaining data from natural experiments may or may not require extensive on-the-ground investigation. Dunning, however, highlights the importance of fieldwork for locating and identifying random assignments that occur naturally in the world (Dunning 2008, 17–19). Scholars acknowledge that these “natural” occurrences are relatively rare, and that extensive knowledge of the precise circumstances that generated the data is necessary to understand whether the assignment is actually random, and thus whether

<sup>32</sup> Interview, BR-17, November 14, 2013.

<sup>33</sup> Ethnographic study could also, they write, explore the emotional reactions of victims who volunteered to participate but were assigned to the control group, perhaps dashing their hopes of confronting the offender and receiving an apology.

the assumptions of experimental research – e.g., that the treatment and control groups are effectively alike in all respects save the treatment itself – are met (Dunning 2012, Sekhon and Titiunik 2012).

In sum, the ways in which other field research techniques augment and facilitate experimental projects, and help scholars to firmly grasp the field context, are many. Regardless of how they are (or are not) combined with other techniques, field experiments entail a series of practical, logistical, and intellectual challenges. The rest of this chapter discusses those challenges, and offers some strategies for addressing them.

## Working with partner organizations

Field experiments, like surveys, can be undertaken by lone researchers but often involve larger research teams and intensive collaboration between academics and partner organizations. Because of the rising demand for scientific evidence on which to base decisions about policy and aid resources, many kinds of organizations have become receptive to partnerships with researchers. Field experiments can involve cooperation with partners of several types, notably governments, intergovernmental organizations, foundations and other private funding organizations, political parties, NGOs, academic research groups, private firms, and social movement organizations. This section explores the kinds of collaborations that political scientists have forged and draws out lessons about maintaining such cooperation.

Different kinds of partner organizations have distinct strengths and weaknesses, and the choice of an appropriate partner will depend on the research question and field context. As Duflo *et al.* observe, partnering with government agencies may offer particularly far-reaching support, and allow research on pilot programs and interventions that have a strong chance of eventually becoming policy. These sorts of partnerships, then, have the potential both to benefit the local context, and to facilitate political science research with real-world implications. Yet such cooperation can be elusive, as it may require the approval of high-level officials and extensive coordination within and among bureaucratic agencies. NGOs, by contrast, may be more nimble and adaptable as research partners than state offices are. Further, as donors increasingly seek compelling evidence that funded programs have payoffs, NGOs often embrace cooperation with researchers who can provide solid program evaluations. It may even be possible to work with for-profit firms (Duflo *et al.* 2006, 20–22).

The collaboration necessary to conduct field experiments with a partner organization has advantages and disadvantages. One of the primary advantages is that sharing expenses with a partner organization can reduce the (often significant) cost of carrying out a field experiment.<sup>34</sup> A partner organization may have human resources, such as staff or volunteers, who can make it possible to apply a treatment at the scale that an experiment requires. Further, the input of partner organizations with deep roots in the field site may be necessary for developing treatments that make sense in context. Local organizations or state agencies can help to legitimate an experimental study and to encourage people to participate in it. In conflict-ridden or other hazardous settings, another possible advantage of partnership is an increase in the security provided for the research team.<sup>35</sup> And, to the extent that the treatment is intended to provide some benefit to the locality or to alleviate a problem, local groups can help ensure that the results of the experiment have a lasting payoff: “so the intervention doesn’t die when I leave,” as one interviewee put it.<sup>36</sup>

Yet partnerships can entail disadvantages and challenges as well. Scholars might have to struggle to balance the policy goals of their partner organization with answering their own theoretical questions (Humphreys and Weinstein 2009; List 2011). In some instances, particularly where partner organizations are not obtaining subjects’ consent prior to applying an intervention, conflicting goals and practices may even present the researcher with ethical dilemmas.<sup>37</sup> Also, researchers and non-academic collaborators frequently face different sets of pressures and incentives. For example, political scientists aim to determine impartially what the data say about the effectiveness of a particular program or intervention. Their counterparts, by contrast, may not be as committed to the integrity of the research process or the reporting of its findings; negative results could, in some cases, jeopardize funding or cast leaders in a bad light. One experimentalist even gave examples of members of partner organizations apparently sabotaging an experimental evaluation of their programs, fearing what the results would show.<sup>38</sup>

<sup>34</sup> Gerber makes this point (2011, 130).

<sup>35</sup> Paluck (2009) explains how working with an NGO allowed her team to travel in well-maintained SUVs in a large convoy and have access to expanded networks of information regarding outbreaks of violence.

<sup>36</sup> Interview, BR-15, October 25, 2013.

<sup>37</sup> Humphreys explores some of the potential ethical challenges facing researchers doing “embedded” experiments with partner organizations (2011).

<sup>38</sup> Interview, BR-18, November 18, 2013.

Sometimes collaborators may not understand the need to follow the experimental protocol precisely. In other instances, groups may simply lack the capacity to carry out what the researcher needs them to do. One scenario, for instance, is that “the [partner] group thinks that we’re going to be able to make this many phone calls, provide this many students, and they grossly overestimate their capacity. If the volunteers don’t show up to do it, there’s not much you can do about it.”<sup>39</sup> People in the partner organization may not be able to administer a treatment in a consistent way to all subjects.<sup>40</sup> For all these reasons, depending on circumstances, it pays to be cautious about delegating crucial elements of the research process – such as random assignment of subjects to treatment groups – to partners, and it is prudent to maintain close, independent oversight of the integrity of the study.<sup>41</sup> Judith Gueron’s essay (2002) deals at length with the tensions that can arise between researchers and partner organizations, which, she writes, often commit to a randomized study without fully understanding its implications and challenges.

More generally, partnerships require substantial start-up investments of time to establish contact and build trust. Training collaborators and coordinating their work also can be time-intensive. Our interviewees described building and maintaining relationships with partners as one of the persistent challenges of their experiment projects. As one described it: “Organizations don’t like to be treated in very transactional ways: ‘You’re a source of data, let me come in, let me have your data!’ They are kind of mistrustful of me: what’s my intention really? Do I really care about the outcome?”<sup>42</sup> Another said that “there’s nothing that keeps me up at night more” than concerns about how to communicate with partner groups, without offending them, in ways that reinforce their obligation to keep to the terms of a memorandum of understanding.<sup>43</sup> One experienced researcher stressed the importance of spending time in person with collaborators, “going out to bars,” building their commitment to the experiment and encouraging them to air, in advance, any concerns about the project or limitations in their capacity to carry it out.<sup>44</sup> Yet the cautions we discuss elsewhere in connection with forming relationships in the field still stand. For instance, some field experimentalists have found that their political neutrality may be perceived as compromised as they become identified as affiliates of the partner organization.

<sup>39</sup> Interview, BR-18, November 18, 2013.

<sup>41</sup> Interview, LM-13, September 7, 2012.

<sup>43</sup> Interview, BR-15, October 25, 2013.

<sup>40</sup> Interview, BR-17, November 14, 2013.

<sup>42</sup> Interview, BR-17, November 14, 2013.

<sup>44</sup> Interview, BR-18, November 18, 2013.

While institutional collaborators may have a lot to offer, it is entirely possible to do field experiments without a partner organization. Researchers working in areas with little NGO activity, and those pursuing theoretical questions that do not happen to dovetail with the agendas of nonprofits or government agencies, all have good reason to go it alone. Political scientists flying solo can apply for grants to support the costs of training their own teams of research assistants to carry out the field experiment. Alternatively, scholars might decide to develop “light-weight” research designs requiring relatively little in terms of institutional support, computer infrastructure, and staff assistance. An example would be an experiment embedded in a survey conducted through face-to-face, in-home interviews, where treatments consist of different question wordings or prompts. One field experimenter explained the choice to refrain from working with an established research institute:

It's still more financially sound to go on my own, especially as I have these connections [from a previous project]. Someday when I have funds, I can completely delegate it. I also still like to be involved in the process, the sampling process, for instance. I don't know how much I'd want to delegate. It's fun too.<sup>45</sup>

---

## Developing and randomizing the treatment

---

Developing the experimental intervention, i.e., the treatment, is a crucial step in the process of experimental research. As noted in the previous sections, ideas for treatments sometimes come from prior phases of an investigator's research, or from collaborators. Often – as with research in other fields, such as clinical trials of drugs – treatments are chosen in part because existing evidence of a causal effect is sufficiently plausible, or related research findings sufficiently robust, to justify the work of putting the posited effect to an experimental test. Treatments may also be driven by programs that partners such as donors, NGOs, or government agencies happen to be implementing and wish to evaluate. Naturally, in political science research, the propositions to be tested also generally follow from hypotheses that investigators aim to test, which in turn relate to broader bodies of theory. Thus, as a practical and intellectual matter, there can be multiple sources underlying the choice of treatments.

<sup>45</sup> Interview, BR-6, August 14, 2012.

Developing a treatment takes substantial time, and draws upon knowledge of the field milieu. Experimenters often use interviews or focus groups to refine the treatment, and partner organizations play an important role in this process. In part, they do so to ensure that the intervention constitutes a proper and fair test of the causal proposition at hand. Additional goals are to be confident that the treatment will not harm subjects in one way or another, even if in intangible ways, such as making them more cynical about politics, less inclined to vote, more hostile toward rival groups, and so forth. As one interviewee put it: “If there aren’t enough other research methods that are applied before the experimental intervention, it can be disastrous, or at least bad. We really owe it to [research subjects] before doing an experimental intervention to check over and over to make sure this isn’t going to go awry.”<sup>46</sup>

Once the treatment has been developed, the researcher must then randomize the assignment of that treatment. As in other types of research design, randomization can occur at multiple levels. In a field experiment on the effect of solar radios on citizens’ perceptions and political views, one scholar identified and matched “twin” villages that were similar in multiple respects, and then randomized which village in each pair received solar radios (the treatment) and which received solar flashlights (the placebo, or control).<sup>47</sup> At the next levels, public lotteries determined which compounds in each village, and which men and women in those compounds, received the items in question.

Yet, particularly in field as opposed to laboratory settings, achieving randomization is not always easy. The following passage, from an essay by Elizabeth Paluck discussing her work in the Democratic Republic of the Congo, nicely illustrates some of the challenges that random assignment may confront in the field:

In an office tucked under palm trees north of Goma, DRC, I squeezed onto a narrow wooden bench with ten researchers and one driver to face the mayor’s desk. “Karibu, welcome,” the mayor smiled at us. I explained that we were evaluating an NGO-produced radio programme about community relations, and I presented him with our *ordres de mission*. He nodded and started to sign and pass them back to us. “Many people will be grateful that an NGO is showing interest in our situation. You are invited to work in the neighbourhood where I live.” I explained that our choice of neighbourhoods and people was random. He smiled a bit regretfully and turned to a faded hand-painted map on the wall. “I should update you on the security situation.

<sup>46</sup> Interview, BR-15, October 25, 2013.

<sup>47</sup> Interview, LM-1, April 13, 2012.

You should not go farther than these neighbourhoods here, because outside there has been some fighting." (Paluck 2009, 38–39)

Whether in an effort to protect the research team or direct resources to his home area, the author's host had something quite different from a random draw of neighborhoods and subjects in mind. As Paluck goes on to explain, serious security threats did impinge on the team's work.

Even when the principal actors involved in a study initially agree on the goal of randomization, many other impediments to random assignment can emerge in field settings. Some obstacles include: inadequate training for implementation or capacity to carry it out; social dynamics among team members; a reemergence of previously reconciled competing goals between a researcher and her partner organization – or a reassessment of the partner organization's priorities; and changes on the ground (e.g., a road closure, grant cancellation, NGO organizational failure, or coup). Several scholars highlighted how the procedures for implementing randomization were often very complex, requiring extensive training and supervision of field enumerators with multiple manuals and tables.<sup>48</sup> Others suggested that partner organizations can balk at implementing designs that require giving something of apparent value to members of a treatment group and withholding it from members of a control group, asking "why can't we just give it to them?"<sup>49</sup>

Indeed, even though randomization is a core *methodological* principle of any field experiment, it is important for researchers to consciously deliberate on the ethical issues involved in randomization in their particular field context. Would randomization of an intervention violate central *ethical* principles of justice and beneficence? For example, simple randomization applied to a public health program might mean that the most impoverished and needy communities were last to receive a vital public health solution. The answers from the literature are not straightforward, and depend, to some degree, on whether scholars are focused on the short-term or long-term implications of the experiment, and whether they consider the process or outcome to be more important. Where Deaton (2009) questions the ethics of randomization, Gerber (2011) turns the challenge on its head, pointing to the long-term costs of *not* carrying out randomized trials and thus failing to obtain systematic evidence about whether interventions actually work.<sup>50</sup>

<sup>48</sup> Interview, LM-13, September 7, 2012.

<sup>49</sup> Interview, BR-15, October 25, 2013.

<sup>50</sup> List (2011) also makes the argument that the cost of not doing randomized experiments outweighs these concerns.

De La O and Wantchekon (2011) argue that, since resources are limited, random assignment of needed and desirable treatments is actually a more transparent and fair decision procedure.<sup>51</sup> But, since individuals or communities often vary in terms of their level and intensity of need, Humphreys and Weinstein (2009) usefully suggest a strategy of randomizing among those with equal need. Another strategy is to include everyone eventually, but randomly assign the roll-out or timing of when the participants receive the treatment (Blattman *et al.* 2013).

From the perspective of the populations under study, seemingly arbitrary assignment of individuals, communities, or other units to treatment groups can cause offense or at least raise questions. Consequently, scholars emphasized the importance of customizing the randomization process for the specific field site and carrying it out in ways that come across as appropriate. In some cases, scholars advocated the use of public lotteries as locally meaningful. For some others, throwing dice was acceptable. For another who worried that this “might look weird” to community members, enumerators used tables from a book.<sup>52</sup>

Finally, scholars should evaluate whether certain treatments can ethically be randomized at all. One interviewee gave an example of a research team that decided against the randomization of a planned treatment because randomizing would have meant preventing legitimate local political authorities from fulfilling their customary roles in the community.<sup>53</sup> Each of these strategies for dealing with potential ethical concerns has costs and benefits for participants and communities as well as the study that can only be assessed with a deep understanding of the specific dynamics of the field context.

---

## Recruiting and preparing subjects

---

Experiments can be conducted in ways that involve minimal contact between researchers and subjects, as in non-field-based studies conducted over the web. Commonly, though, a considerable part of the work involved in setting up an experiment lies in recruiting, preparing, and debriefing the people who participate in the study. In some experiments, participants are randomly

<sup>51</sup> A similar argument about the greater fairness of randomization was made in several interviews.

<sup>52</sup> Interviews, LM-1, April 13, 2012; LM-13, September 7, 2012.

<sup>53</sup> Interview, LM-13, September 7, 2012.

selected from particular populations, and in others they are recruited through other means. In order to avoid violations of protocol or invalidation of the experiment, all subjects within a given arm of a study must be given the same instructions and guidance, and be kept from influencing one another in the course of the experiment. As one researcher observed, providing just a single example of precautions that must be taken: “You have to explain to them why they can’t first watch their friend do the experiment before they do it themselves.”<sup>54</sup> In this section, we review a range of considerations involved in recruiting subjects and preparing them to participate in an experiment.

When recruiting participants, the researcher must carefully consider where to do so; how many participants to recruit; and whether and how to compensate subjects for participating in the experiment. Ideally, a scholar’s research questions shape the choice of appropriate subject populations. Yet, as we noted previously, in many field experiments researchers collaborate with organizations that may have already identified certain countries, regions, communities, or demographic groups that they wish to target with their intervention.

The question of how many subjects to recruit is related to the issue of how many hypothesized explanatory factors there are to test. While different research questions in different field contexts may necessitate the inclusion of more or fewer, simplicity is usually advised. Proliferating explanatory factors or conditions complicates the analysis of interaction effects, multiplies the number of treatment groups needed, and thus increases the number of subjects who must be recruited in order to obtain the same statistical power.<sup>55</sup> Knowledge of the field context is critical for anticipating whether it is logistically feasible to recruit the desired number of subjects in the available time period.

The ideal number of subjects is also tightly connected to the issue of compensation. Normative standards on subject compensation differ widely by discipline. Psychologists rarely pay subjects on the basis of specific actions they take during an experiment; economists often pay participants contingent on their performance in the experiment; and political scientists seem to be split between these practices (Druckman *et al.* 2011b, 7). If monetary incentives are used in a field experiment, the researcher must devise an appropriate scale of incentives. Morton and Williams (2010) attempt to

<sup>54</sup> Personal communication, June 19, 2011.

<sup>55</sup> Statistical power refers to the probability that a test will reject a false null hypothesis.

summarize existing disciplinary norms with an estimate that participants are paid 50 to 100 percent above the minimum wage for the time spent in the lab. Of course, the opportunity cost of time may be more readily estimated when the experiment takes place in a lab, and not the field. Field settings also vary considerably in terms of the normative expectations for immediate financial or in-kind compensation versus a more generalized commitment of long-term reciprocity.<sup>56</sup>

Once participants have been recruited and the appropriate level of compensation established, the next issue to consider is the kind of preparation they will receive. In many field experiments, participants may not know they are involved in an experimental study. Or, they may know that they are engaged in a research project but be unaware to some degree about the central purpose of the investigation. This aspect of field experiments raises the question of how much deception is ethically acceptable. While deception (or incomplete disclosure) may amplify the realism and hence the internal validity of a field experiment, its potential lasting implications for participants and social outcomes raise ethical concerns. For example, some political scientists have studied public officials' responses to contact from citizens in ways that involve deception. In such studies, emails have been sent that appear to be from real voters or constituents, when actually the sender was a fictitious identity created by the researchers, for instance with the purpose of testing differential responses to constituents from varying racial or political backgrounds. Researchers behind such studies have defended them, arguing that deception is necessary to study the phenomenon in question; that the request is routine and the imposition is modest; and that public officials deserve and expect special scrutiny (Butler and Broockman 2011; McClendon 2012).<sup>57</sup> As with compensation, the use of deception varies by discipline so no easy consensus exists. Where psychologists frequently incorporate deception, economists almost never do, and political scientists are fairly divided on the issue (Morton and Williams 2010; Dickson 2011, 65–67; Druckman *et al.* 2011b),<sup>58</sup> often depending on the source of theoretical and methodological inspiration for

<sup>56</sup> We offer additional thoughts on compensation in Chapter 4.

<sup>57</sup> Another question, also addressed by McClendon's article, is to what extent such deception might have adverse effects on future research by leaving subjects, or others who learn about the study, with negative feelings.

<sup>58</sup> Druckman *et al.* (2011b, 7) find that 31 percent of laboratory experiments published in the APSR through 2005 used deception. To our knowledge, no similar assessment has been done with field experiments.

the study and involvement in multidisciplinary collaboration. We simply suggest that scholars critically reflect in advance about what seems methodologically necessary *and* ethically appropriate given the political and cultural context of their field sites.

---

## Applying the treatment and following up

---

Actually conducting the experiment and applying the treatment pose a set of challenges that are often quite specific to experimental research, although principles and strategies for coping with them overlap with those that apply to other methods. Just as with survey research, patient and repeated pre-testing is essential. Everyone on the research team should be well-practiced in the procedures for applying the treatment and for handling contingencies. Investigators go to great lengths to try to standardize the experience for all subjects, minimizing or nullifying differences except for those assigned as part of the treatment. For example, enumerators or research assistants who have personal contact with subjects may be randomly assigned to avoid bias stemming from systematic effects of their particular backgrounds. They are also carefully trained in advance regarding how to respond to queries from subjects, requests to withdraw from the experiment, or other special circumstances.

The need to keep subjects in the treatment and control groups independent from one another is vital in experimental research, and this creates a distinctive set of challenges for field investigators. The problem of members of one group being influenced by the experimental conditions to which members of another group are assigned is known as interference (Gerber and Green 2012, 8). One form of interference is spillover. Spillover occurs when “those who are treated in turn alter their behavior in a way that affects other subjects” (Gerber 2011, 130). It presents a problem because control groups experiencing the effects of the treatment in some form or another may bias estimates of the treatment’s actual effects.<sup>59</sup> As Gerber states, while laboratory experiments generally are not subject to this problem, it is more likely to crop up in field experiments. An example would be a project in which certain state senators were randomly assigned to a treatment (for instance, being told that details of their campaign finances would be

<sup>59</sup> Gerber and Green’s book contains a chapter on interference among experimental units (2012, Chapter 8). See also Gerring (2012, 242–246).

publicized to constituents) while others were randomly assigned to a control group. If, during the course of the experiment, senators in the former group changed their behavior in response to the treatment (turning down questionable campaign contributions, for example) and their colleagues in the control group emulated them, this would be a clear case of spillover.

The potential for spillover varies with features of the research design, the treatment, and the research setting. One experimentalist we interviewed struggled with the problem of keeping treatment and control groups separate from one another when the two groups were part of the same organization and had the same leader, and the experiment was to last for weeks. By contrast, maintaining the independence of treatment and control was much easier in studies where subjects were given the treatment individually, whether by email or by knocking on their doors, and outcomes took the form of the subjects' immediate responses. The interviewee observed: "This is kind of an ongoing limitation of field experiments: it's easier to study things where the treatment is really short, but there are a lot of things in the political world where the treatment takes a long time."<sup>60</sup>

Familiarity with the social environment of the field site helps investigators to determine how serious a problem spillover might be and what measures might be taken to prevent it. Evaluating the potential for spillover may require an in-person visit to determine whether households or neighborhoods assigned to treatment are physically proximate to, or are otherwise likely to interact with, those assigned to the control group, for instance.<sup>61</sup> One obvious approach to preventing spillover is to design an experiment in such a way that treatment and control groups are geographically separated. For example, in a study of voter responses to information about candidates' corruption, de Figueiredo, Hidalgo, and Kasahara selected 400 voting precincts in São Paulo through an algorithm that maximized the distance between the treatment and control groups.<sup>62</sup> After lamenting that there were "so many opportunities for spillovers to happen," another scholar advised frequent and open communication with partners and the research team, concluding: "If you are on the ball and can catch them [spillovers] quickly,

<sup>60</sup> Interview, BR-17, November 14, 2013. <sup>61</sup> Interview, BR-2, July 30, 2012.

<sup>62</sup> "When do voters punish corrupt politicians? Experimental evidence from Brazil," unpublished paper, p. 15. The algorithm also incorporated other information that the authors acquired through careful investigation of the city's neighborhoods, obtaining relatively even distributions on income levels, and vote choice in previous elections. As well, the algorithm avoided high-rise neighborhoods in which it would be difficult to ensure that the treatment (flyers with information about candidates) reached the mailboxes of individual households.

then you can save the project.”<sup>63</sup> Another approach is to incorporate spill-over into research designs deliberately, measuring it and analyzing its effects. Nickerson, for example, had canvassers in two American cities knock on doors and deliver a get-out-the-vote appeal prior to the 2002 Congressional primary elections. He measured the effect of this exhortation on the propensity to vote of both the person who answered the door and on a second household resident who did not receive the appeal, thus measuring how civic participation is spread from one person to another (Nickerson 2008).

Finally, in some experiments, the last interactive step is debriefing participants afterward to solicit additional information from them, as well as to give them further information about the study in which they took part. Investigators may, for instance, send correspondence to subjects; they may request that they participate in a follow-up survey; they may hold community feedback meetings; or they may sit down with them for open-ended interviews. This is another aspect of experimental research that can be simple in laboratory settings but more difficult in the field. Indeed, in many field experiments, particularly those in which subjects are unaware that they are part of an experiment, no debriefing is conducted.

One purpose of debriefing is to reveal problems that may have occurred in the process of the experiment, such as possible non-compliance on the part of researchers (or their partners) or participants, resulting in a failure to treat. One source of non-compliance is that certain randomly selected individuals and communities might be more difficult to access than others, producing a systematic bias in the treatment group. Another source of non-compliance is individuals’ inability or unwillingness to follow instructions. In this instance, the experiment successfully reached the participant, but the person did not actually read, do, or complete what was intended to be the treatment. A second type of problem that debriefing can reveal is that the intended treatment was confounded with something else. In this case, the treatment was designed to create a certain kind of stimulus, X, but instead was perceived as something else, X<sub>2</sub>. Following up with participants can also help scholars to interpret the findings from the experiment, for instance by shedding light on the mechanisms through which the treatment exercised its causal effect, or, if no average treatment effect was found, on the reasons why that was so. A further goal can be to determine how enduring any outcomes of the treatment are.

<sup>63</sup> Interview, LM-13, September 7, 2012.

The purpose of following up with participants after an experiment is not merely to provide experimenters with further data and to detect problems in the execution of the study, however. Debriefing can also play important roles in conveying information to participants about the research they took part in. At a minimum, IRB guidelines suggest that subjects should be debriefed in projects “involving deception or incomplete disclosure, especially if the research may induce psychological stress, guilt, or embarrassment,” allowing the investigators to “explain any deception involved and to help the subjects deal with any distress occasioned by the research.”<sup>64</sup> Some experimentalists expressed a belief that investigators have even broader obligations to explain their research to those involved. In part, this is a matter of clearing up puzzlement and suspicion. As one said: “It’s kind of ridiculous to not debrief. I’ve seen lots of cases where the subjects of the field experiment are aware that something totally unnatural and weird happened,” and they deserve to have this explained. This interviewee further stated that

disclosure or debriefing is important because I believe that the research should be useful to the people who participated in it, and not just in a trickle-down way, in that it will be useful to them because I find an effect and then convince the World Bank or the government [to act on it], but that this thing that we did is useful in everyday life.<sup>65</sup>

According to this view, follow-up should convey research findings and also engage participants in two-way deliberation about what was learned.

---

## **Conclusion**

---

Field experiments, as well as other types of experiments, are becoming more and more prevalent in the study of politics. Clearly, an increasing amount of experimental work is being conducted in the field, or is working its way through the publishing pipeline. Whether or not field experiments revolutionize social policy, as high-profile advocates like Esther Duflo have claimed they will, these methods are being applied to pressing problems around the world and are yielding new insights. Experiments sharply contrast with the other data-collection techniques considered in this book in certain respects.

<sup>64</sup> Institutional Review Board Guidebook, Chapter III, accessed December 5, 2013 at [www.hhs.gov/ohrp/archive/irb/irb\\_chapter3.htm](http://www.hhs.gov/ohrp/archive/irb/irb_chapter3.htm).

<sup>65</sup> Interview, BR-19, December 4, 2013.

To conduct an experiment requires a narrowly focused research question and well-defined causal factors and outcome measures. The random assignment of subjects to treatment groups and the deliberate introduction of treatments has no parallel in methods like in-depth interviews and archival research.

And yet a clear theme running throughout this chapter is that effectively conducting field experiments requires deep engagement in a research context and the use of multiple methodologies and skills. The processes and practical challenges of carrying out such research have a great deal of overlap with, and often entail, other forms of fieldwork. This theme resonated in all of our interviews with veteran experimentalists, and of course, it reinforces the core principles articulated in this book. Each stage of experimental research – from conceiving the idea of a treatment to assess, to working with partner organizations, to designing the experimental protocol, to implementing it, to interpreting the results – can benefit from a degree of immersion in the field setting and the use of qualitative techniques such as interviews and site-intensive methods. Still more field research is required as experiment-based research programs are replicated and expanded in new environments and with new types of subject populations, in order to understand how local contexts vary in theoretically relevant ways and to adapt treatments, measures, and protocols appropriately.

As with other data-collection techniques discussed in this book, this chapter also showcases the value of making the methodology of field experiments explicit and transparent in the publication of articles and books, or in standalone essays. Practically every experimental research project that we discussed with our interviewees had an extensive back-story, often with fascinating twists and turns. Such lore is shared informally within the experiment community – yet disseminating it more widely could help a broader population of scholars learn what it takes to conduct a field experiment, thus informing future research. We contend that such transparency, which would shed further light on the strengths and weaknesses of these promising techniques,<sup>66</sup> is particularly important for field experiments, since they are relatively new in political science. Broadening methodological debates around experimental methods could, in turn, stimulate an expansion of the types of research questions experiments address, and of the types of experimental design employed.

Experimentalists sometimes object to the idea that their technique raises special ethical considerations, feeling that they are being singled out unfairly.

<sup>66</sup> Interview, LM-9, August 30, 2012.

There is some validity to this. Yet it is also true that most ethical questions stem precisely from those aspects of field experiments that make them distinct and intellectually appealing. Field experiments isolate the effects of a potential causal factor by bringing it into existence in the form of a treatment, and selectively applying it across groups. Intervening in the world in this way helps make field experiments particularly relevant to policy issues and the practice of politics generally – and also necessitates particular forms of scrutiny. The community of experimentalists within political science is in the early stages of addressing some of the ramifications of this research. Merely adhering to the letter of IRB requirements may not always be adequate to avoid harm, let alone to meet higher standards of beneficence. Much thought and consideration are needed as the discipline works out what it means to practice ethical commitment in this mode of research.