5 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.[1] 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,

[1] 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.[2] 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.[3]

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

[2] 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.

[3] 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,[4] 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.

[4] In some of these projects, scholars may have been building collections of new data rather than collecting pre-existing data.

https://doi.org/10.1017/CBO9780511794551.006 Published online by Cambridge University Press

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.[5]

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.[6] 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

[5] Tosh (2000, 27–47) outlines and discusses, from a historian's viewpoint, the main categories of documentary material.

[6] 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.

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

Source materials	Locations
• Posters, postcards, bumper stickers, graffiti, invitations, flyers, popular-culture visual or audio materials (printed cloth, art, songs, etc.), and other informal/unofficial materials	The broad research setting
• Maps, satellite imagery, geographic data • Official documents, files, reports (diplomatic, public policy, propaganda, etc.) • Government statistics (e.g., on population, sewage systems, roads, GDP, budgets, immigration, trade, poverty) • Correspondence, memoranda, communiqués, queries, complaints • Parliamentary proceedings • Minutes from intra-governmental and inter-governmental meetings • Testimony in public hearings • Speeches, press conferences • Military records • Court records • Legal documents (charts, wills, contracts) • Chronicles, autobiographies, memoirs, travel logs, diaries	Archives (national or local, government or private) Government agencies/entities (e.g., congressional offices, ministries, etc., at various levels)
• Private papers • Brochures, posters, flyers • Press releases, newsletters, annual reports • Records, papers, directories • Internal memos, reports, meeting minutes • Position or advocacy papers, mission statements • Party platforms	Individuals or organizations (e.g., political parties, unions, businesses, hospitals, schools, religious entities, interest groups, universities, NGOs, etc.)
• Radio broadcasts (transcripts) • TV programs (transcripts) • Magazines, newspapers • Electronic media	Media
• Published collections of documents, gazetteers, yearbooks, etc. • Private papers • Books, articles, dissertations, working papers	Libraries (university, national, local)

https://doi.org/10.1017/CBO9780511794551.006 Published online by Cambridge University Press

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.[7] 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).[8] 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 purposively. 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.[9] 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.[10] 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

[7] Interview, DK-1, July 20, 2012.

[8] 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.

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

[10] 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.[11] 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.[12] 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.[13] 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

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

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

[13] 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.[14] 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

[14] Chapter 4 offers examples, and Chapter 11 discusses how new technologies are shaping the future of field research.

https://doi.org/10.1017/CBO9780511794551.006 Published online by Cambridge University Press

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.[15] 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.[16]

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.[17] 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.[18] 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.

[15] Interview, DK-3, July 27, 2012.
[16] Interview, LM-25, October 21, 2013.
[17] Interview, DK-5, July 31, 2012.
[18] 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

https://doi.org/10.1017/CBO9780511794551.006 Published online by Cambridge University Press

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.[19] 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.[20]

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.[21] 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

[19] Example mentioned by a participant in a workshop for this book manuscript held at Indiana University, December 2, 2011.
[20] Personal communication, November, 26, 2010.
[21] 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.[22]

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.[23] 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"

[22] Interview, DK-9, August 1, 2012. [23] Personal correspondence, January 30, 2011.

https://doi.org/10.1017/CBO9780511794551.006 Published online by Cambridge University Press

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.[24] 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.[25]

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.[26] 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.[27] 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.

[24] Personal communication, James Mahoney, March 23, 2011. [25] Interview, DK-2, July 26, 2012.

[26] 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.

[27] 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!"[28] 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.[29] Government ministries, or any

[28] Interview, LM-15, September 10, 2012. [29] Interview, DK-5, July 31, 2012.

https://doi.org/10.1017/CBO9780511794551.006 Published online by Cambridge University Press

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.[30] 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.[31] That is, they will need a selection or sampling strategy – a plan for choosing a subset of the available materials.[32] 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?

[30] Personal communication, November 20, 2010.

[31] 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.

[32] 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).[33] 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.[34]

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.[35] 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.[36] On the other hand, even though most scholars will select sources purposively, random sampling from a wider

[33] Trachtenberg's Appendix II (2006, 217–255) identifies an extensive range of published, semi-published, online, archival, and open sources relevant to international history.
[34] Interview, BR-4, August 9, 2012.
[35] Elizabeth Saunders, 2010, "Archival methods of research," IQMR, Syracuse University, Syracuse, NY.
[36] 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.[37] 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.[38]

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.[39] 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

[37] 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.

[38] See Stockmann (2010, 113–116) on drawing samples from media sources.

[39] Interview, James Mahoney, March 23, 2011.

reports (including statistical data) rather than delving into a plethora of fine-grained documents from multiple ministries.[40]

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.[41] To briefly reiterate our discussion from Chapter 4, scholars might take notes, take digital pictures,[42] photocopy,[43] 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.[44] 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*

[40] Personal communication, junior comparativist, November 26, 2010.

[41] 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).

[42] 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.

[43] See Hill (1993, 53) for a compelling set of advantages to photocopies.

[44] 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.[45] 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.

[45] 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?[46] 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.[47] 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

[46] 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.

[47] 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.[48] 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.[49] 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.[50] The available sample (particularly with regard to historical sources) is likely not random, and may well be unrepresentative of the broader whole, and/or

[48] Interview, DK-6, July 24, 2012.

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

[50] 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).[51] 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."[52]

Different types of pre-existing sources may have different biases.[53] 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

[51] 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).

[52] IQMR at Syracuse University in June 2010.

[53] 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.[54]

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.[55] 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,[56] and evaluate their evidentiary value.[57] Those who take an interpretive approach to analysis

[54] 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.

[55] The systematic exclusion of certain voices is one rationale for conducting oral history interviews, discussed in Chapter 6.

[56] 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.

[57] 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?[58] 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,

[58] 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)?[59] Is the version accessed the only version there is? Was it published?[60]

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.

[59] Dibble (1963, 204–210) discusses criteria and rules for evaluating testimony (produced by individuals) and "social bookkeeping" (produced by groups or organizations).

[60] 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.[61] 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

[61] 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.[62] 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).[63] 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[64] 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

[62] Parts of this section draw on modules taught by James Goldgeier and Elizabeth Saunders at IQMR, and presentations at George Washington University's SICAR.

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

[64] 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.[65] 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,

[65] 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.

https://doi.org/10.1017/CBO9780511794551.006 Published online by Cambridge University Press

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.[66] 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."[67]

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.[68] 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.[69] Researchers can also attend institutes and workshops,[70] 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

[66] Interview, DK-2, July 26, 2012. [67] Interview, BR-4, August 9, 2012.

[68] 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.

[69] 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.

[70] 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,"[71] H-Net Discussion Networks in general (and H-Diplo Resources in particular, especially "Guides to doing research" under "Organizations"),[72] and the American Historical Association's web site.[73] 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).[74]

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);[75] their social practices and informal rules (for instance, dress code);[76] and what can be brought into the archive (technology in particular).[77] 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?[78]

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

[71] www.archivesmadeeasy.org. [72] www.h-net.org/~diplo/resources.
[73] www.historians.org/info/research.cfm. [74] Interview, DK-5, July 31, 2012.
[75] Personal communication, January 30, 2011.
[76] In particular, they might check whether the use of ultra-thin gloves to protect the documents is required.
[77] 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).
[78] 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.[79]

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.[80] 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

[79] Hill (1993, 27–39) discusses name-oriented and other archival search strategies.

[80] 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.[81] 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.[82]

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,

[81] 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.

[82] 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.[83] 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.[84] 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.[85] 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.[86] 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

[83] Interview, LM-7, September 20, 2012. [84] Interview, DK-17, August 24, 2012.

[85] 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."

[86] 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.[87] 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.[88] Further, a system for highlighting items that look potentially useful (or less useful) and justifying

[87] 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.

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

https://doi.org/10.1017/CBO9780511794551.006 Published online by Cambridge University Press

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,[89] 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.

[89] 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.

https://doi.org/10.1017/CBO9780511794551.006 Published online by Cambridge University Press

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.[90]

[90] 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.