HOW TO REPORT INTERVIEW DATA

Erik Bleich and Robert Pekkanen

Interviews offer a potentially valuable yet often neglected or mistrusted source of evidence in the investigation of political phenomena. Readers may fear that the interviews were gathered from unrepresentative individuals and are thus biased, that the information gained through the interviews is inaccurate, or that the interviewer may distort the evidence through selective presentation undetectable by the readers. We believe that much of this mistrust can be alleviated if scholars follow the procedures we recommend below for reporting how interview evidence was gathered. We see these procedures as extremely helpful not only for researchers that rely primarily on interview data, but also for those—most likely the majority of our readers—who deploy interviews as one type of evidence in a multi-method research project. Our aim in this chapter is to create guidelines that will increase the rigor and transparency of the interviewing process and therefore enhance readers’ confidence in interview data. This will encourage scholars to use interview evidence more frequently and systematically, enabling gains in knowledge by unsheltering this stream of evidence.

Such a goal would be laudable at any time, but it is especially in tune with the recent burst of interest in qualitative methods. Over the past several years, scholars have reinvigorated the use of qualitative research, arguing that it has advantages over quantitative methods in many contexts. This renaissance is not just about redeeming qualitative methods, but also about improving them. It is perhaps inevitable that the uses of interview evidence should enter into discussion of this research program. This volume itself reflects this development, as does work on qualitative data archiving spearheaded by Elman, Kapiszewski, and Vinuata (2010) and related arguments about standards for active citation (Moravcsik 2010).

After touching on the uses made of interview data in various research projects, we identify three key problems plaguing the reliability of interviews as a source of information in social science projects. We then suggest several concrete solutions to each of these problems. Finally, we detail how researchers can formulate an interview methods appendix and an interview methods table as tools that help communicate the reliability of information garnered through interviews.

The Uses of Interview Data

In establishing this connection between the new qualitative methods literature and interviews, Tansey (2007) compellingly argues that interviews make particularly good evidence for process-tracing research. However, we contend that interviews as a stream of evidence are generally compatible with most research goals. For example, Lynch (this volume, chapter 1) details the uses of interviews for preliminary research, the main focus of the research, part of multi-method research, as well as for “window dressing.”

All too often, interviews are utilized only in this tepid final category. There is no denying that a vivid quotation can enliven a scholarly article or memorably summarize an argument. We certainly do not advocate excising these with the effect of making scholarship less readable. However, we feel that scholars use interviews disproportionately as illustrations, not as evidence, because authors and readers alike are concerned that interview data may be less reliable than quantitative data or than written primary or secondary sources. The limitations of these other sources, and the problems of quantitative data reliability in particular, are widely acknowledged. Therefore, there is no reason that interviews cannot form a valuable, if also imperfect, source of systematic evidence for political scientists. However, quoting from a single interview as an illustration while rigorously presenting statistical analyses does not equate to relying on both sources equally.

To highlight the parallels between different forms of evidence, it is useful to consider briefly the differences between surveys and interviews. In contrast to
interviews, surveys are widely deployed as evidence not only by political scientists, but also by scholars from a variety of disciplines. Like interviews, surveys rely on information and responses gained from human informants. Gathering and assessing survey data involves many well-understood complications. Surveys respond to these challenges by reporting their methods in a manner that enables others to judge how much faith to place in the results. We believe that if similar criteria for reporting interview data were established, then interviews would join surveys as a more widely trusted source of evidence. After all, surveys can be thought of as a collection of short (and sometimes not so short) interviews. Surveys and interviews thus fall on a continuum, with trade-offs between large-n and small-n studies. Just as scholars stress the value of both types of studies depending on the goal of the researchers (Lieberman 2005), both highly structured survey research and semi- or unstructured small-n interviews, such as elite interviews, should have their place in the rigorous scholar’s tool kit.\(^3\)

**Problems with Interview Data**

We believe that the pervasive skittishness about the reliability of interview data stems from three fundamental challenges: representativeness of sample; type and quality of information obtained; and accuracy of reporting.

**Representativeness of Sample**

Authors who have conducted numerous interviews typically report some information about their interlocutors. They usually provide names and job titles, or mention the places and dates of the interviews, or both. Whenever an interviewee is quoted, this information appears as a citation, as is appropriate. At times, authors also summarize their interviews as a body. They may state that they interviewed sixteen NGO leaders, eight bureaucrats, and three local politicians. This kind of reporting serves the dual purpose of giving the reader a sense of the author’s sources while also, perhaps, impressing the reader with the author’s authoritativeness on the subject.

Such reporting, however, does not go far enough. This kind of reporting leaves us in the dark about nonresponse bias. Surveys always report response rates, because the higher the response rate, the more valid the survey results are generally perceived to be. Such nonresponse bias might also skew results in the interviewing process. In a set of interviews about attitudes toward the government, for example, those who decline to participate in the survey might do so because of a trait that would lead them to give a particular type of response to the interviewer’s questions, either systematically positive or negative. If so, then we would be drawing inferences from our conducted interviews that would be inaccurate, because we would be excluding a set of interviewees that mattered a great deal to the reliability of our findings.

Currently, we have no way to assess response rates or possible nonresponse bias in interviews; the standard process involves reporting who was interviewed, but not whom the author failed to reach, or who declined an interview. Even more troubling, some common sampling techniques can easily exacerbate biased results. For example, the “snowball” technique refers to the process of seeking additional interview leads from one’s interviewees (also see Martin, this volume). In an environment where interviews can be hard to come by (say, United States senators), the technique has an obvious attraction. Important actors approached with a referral in hand are more likely to agree to an interview request than those targeted through “cold calls.” In addition, if the original interviewee was a good source, then she is likely to refer the researcher to another knowledgeable person. All in all, snowball sampling has much to commend it as a technique.

Yet, one danger with this strategy is that researchers become trapped within a network of interlinked respondents who see the world through the same lens. Snowball sampling clearly introduces the possibility of bias, because the original interviewee and subsequent contacts may share similar views on the subject of the interview.\(^4\) This could involve deliberate manipulation by sophisticated political actors, but it need not. For example, in researching a controversial policy decision, an author may be steered toward interviewees who all agree that the right decision was made, simply because those are the only people that the initial interviewee (and the source of the snowball referrals) feels are “worth” talking to—because anyone who really understands the issue would have to agree with the decision, and so anyone who disagrees obviously does not understand it and is not worth talking to. After interacting with a number of bureaucrats, politicians, and interest groups linked through the snowball technique, the researcher could conclude that she has reached a well-rounded consensus view on the issue at hand. In actual fieldwork, researchers can mitigate these problems by selecting initial interviewees from a wide variety of backgrounds and by being attentive to the perils of relying too heavily on information obtained from one person’s referrals. An important point for us is that when snowball sampling is used to generate some or all interviews, and the “snowballed” interviews are not indicated in any way, readers are left without valuable information needed to enable them to assess the reliability of the interview data.
Type and Quality of Information Obtained

The second problem is that the interviews might not have produced accurate information, regardless of how many or what variety were conducted. This could happen for a variety of reasons. Interviewees might not transmit accurate information, perhaps because they are ill informed or because they are intentionally deceiving the interviewer to bolster their reputation or for other purposes. For example, one of the authors of this article conducted an interview with a former prime minister upon referral by a senior journalist. The author opened his questions with a query about the role of factions within the prime minister's party, to which the politician replied that factions simply did not exist. It may have been the case that the senior journalist's presence in the interview inhibited this politician from speaking frankly, or perhaps he simply could not be bothered to tell the truth (it is extremely unlikely that he believed what he said). This meant that the interview data were of extremely low quality, and in fact the author chose not to incorporate the data in the results of his research. In other cases, of course, interviewees say things that are quite surprising, against their own interests, and very credible. Both the authors have been told things in interviews that would, if made public, result in the firing of the speaker, or in the near certain loss of office by the politician who confided in the author.

On the other side of the equation, the interviewer might not be skillful enough to ask the right questions or to understand what was being communicated, might deliberately misinterpret the information communicated and thus falsely use it as evidence, or might unintentionally misconstrue the interview data, perhaps because of subconscious biases. In one author's experience, his interpretation of a key interview differed dramatically from that of a second scholar who was also present. After lengthy discussion it became apparent that the other scholar had not fully grasped the subtleties of the statements because of imperfect fluency in the language in which the interview was conducted. The point here is not to highlight linguistic competence, but rather the perils of inaccurate reporting. If the other academic had cited these interviews, suitably anonymized, in a publication, readers would have no way of knowing that this account was based on a complete misunderstanding of the interview.

These types of challenges give rise to well-founded concerns. Readers may harbor serious doubts about the quality and reliability of data gleaned from interviews. Just as with the problems associated with quantitative studies, these misgivings can never be totally overcome. Yet, just as in quantitative studies, we can think of them as the "measurement error" of the process of collecting information or data through interviews. Viewed in this light, they are not insurmountable problems, but rather garden-variety challenges to which scholars must develop the most compelling solutions possible.

Accuracy of Reporting

The third problem arises when authors selectively report interview data in a way that generates or reflects a particular bias. Anyone who has ever read a full movie review, then seen a few glowing words quoted out of context on an advertisement for that movie, is aware of this phenomenon. The interviewer could choose the few quotations that illustrate her points, using the evidence to support her argument. Processes of cognition also influence researchers to weigh interpretations provided at the beginning of the interview process more heavily as compared to those provided at the end, and to seek out and convey evidence in support of their argument over disconfirming evidence.

Along with the problems of gauging the quality of data obtained, there are thus parallel problems with assessing what data (out of all data gathered) is reported. Without access to the complete set of interview data, we cannot tell if the spicy quote from the Italian Foreign Ministry official represents the consensus opinion in the government, or is the ranting of a lone crackpot which happens to be well suited to illustrate the author's arguments. Because of the tendency to credit quotes that agree with our own understandings of the situation, this kind of selective reporting is likely to creep into the research even of conscientious scholars. Because there is typically no way for the reader to know the full content of the interviews, it is virtually impossible to judge the accuracy or representativeness of the statements conveyed by the researcher.

These three problems shortchange us all, because they limit the potential contribution to our understanding of political phenomena that may be derived from interview data. When we lack important information about the evidence, we cannot have adequate faith in interview data.

Solutions

Qualitative social scientists can benefit from a common set of standards for reporting the reliability of their data so that readers and reviewers can judge the value of their evidence. As with quantitative work, it is impossible for qualitative researchers to achieve complete reliability. But producers and consumers of qualitative scholarship profit from being more conscious about the methodology of interviewing and from being explicit about reporting uncertainty. This increases the value of interviews in social scientific research, elevating them
from the status of supplementary information or simple adornments to that of widely accepted evidence that can contribute to developing and testing causal theories.

Representativeness of Sample

For any research that relies on interview data as a significant component of theory development or testing, it is important to move toward systematic sampling (Lynch, chapter 1). For some projects, especially those that attempt to gauge the general views of a broad population, systematic sampling will entail random samples of business leaders, bureaucrats, or politicians (Aberbach, Putnam, and Rockman 1981; Martin and Swank 2004). In the types of interviews we focus on, generally related to particular policy or political decisions, there is typically a narrower population of relevant actors, and random sampling is not likely to be the most appropriate or efficient methodology.  

Yet it is often possible for the researcher to identify a theoretically motivated set of target interviewees prior to going into the field. Doing this in advance of the interviews, and then reporting interviews successfully obtained, those refused, and those where the interviewee never responded, has many benefits. For one, this kind of self-conscious attention to the sample frame allows researchers to hone their research design before they enter the field. After identifying the relevant population of actors involved in a process, researchers can focus on the different classes of actors within the general population—such as politicians, their aides, civil servants from all relevant bureaucracies, NGOs, knowledgeable scholars and journalists, and different types within the classes—progressive and conservative politicians, umbrella and activist NGOs, for instance. Drawing on all classes and types of actors relevant to the research project helps ensure that researchers receive balanced information from a wide variety of perspectives. When researchers populate a sample frame from a list created by others, the source should be reported—whether that list is of sitting parliamentarians or business leaders (perhaps drawn from a professional association membership, as Martin describes in chapter 5).

Reporting the sample frame is a vital first step, but it is equally important to report the number of interviews sought within the sample frame, the number obtained, and the number declined or unavailable. This process allows readers to better assess the evidence the author has gathered. Knowing that the author set out to interview ten people, including three politicians, three bureaucrats, and four NGO leaders, but only succeeded in talking to three NGO leaders and one bureaucrat will suggest that a certain type of information may be systematically absent from the analysis. It also gives readers more confidence in the evidence that is presented. Reporting the nonresponse rate encourages researchers to think carefully about bias and to devise strategies to compensate for potential biases. It also allows readers to assess the researchers’ strategies.

Of course, any experienced field researcher knows that the ideal list of interviewees can change dramatically in the field. For example, a rival political party can be discerned working actively behind the scenes, and the field researcher from the example above now needs to add two more politician interviews. Or, one of the bureaucrats interviewed advises that the key person within the bureaucracy is not on the interviewees list. Researchers should always ask interviewees for recommendations for additional interview subjects. This snowball sampling technique can effectively reveal networks or key actors previously unknown to the researcher, thereby expanding the sample frame. We do not advocate eschewing these interviews because they were not on the original list. But we do argue that the researcher should report these developments to readers and expand the sample frame accordingly.

What is most useful, however, is not simply reporting that the researcher engaged in snowball sampling. Rather, the crucial element is reaching the point of saturation in the interview process. At saturation, each new interview within and across networks reveals no new information about a political or policymaking process (Guest, Bunce, and Johnson 2006). If respondents are restating the same causal process as previous interviewees, if there is agreement across networks (or predictable disagreement), and if recommendations for further interviewees mirror the list of people the researcher has already interviewed, then researchers have reached the point of saturation. Researchers must report that they reached saturation to convey to readers that they have exhausted the relevant information to be gained from interviews. Over time, reporting sample frames and interview results will allow scholars to learn from each other's methods (and not just from each other's findings) and arrive at superior interview and qualitative methodology.

Type and Quality of Information Obtained

In our experience, the type and quality of information obtained in interviews can vary dramatically. Some "interviews" consist of two minutes of an influential politician's walking time between meetings, which reveals little or no useful information. Other interviews may consist of multiple conversations with key bureaucrats during which they take the interviewer through the intricacies of complex policy decisions and disclose previously unpublished details that cast the decision-making process in an entirely new light. Some interviews are on the record, others are off the record, and still others are on background. Sometimes
interviewers use recording devices or take notes while the interview is in progress; other times they run to the nearest park bench after the event to empty the contents of their memories into their notebooks.

Researchers also deploy a plethora of strategies for obtaining information from their subjects. They may conduct structured, semi-structured, or unstructured interviews, which range from asking each interviewee a predetermined and consistent set of questions, to simply letting the conversation go where the interviewee takes it. They may obtain interviews through the equivalent of academic cold calls or by ingratiating themselves with the interviewees' colleagues or even family members. They may ask key questions out of the blue to get an honest and unguarded reaction, or they may prime the respondent by giving the context for the question and perhaps even some other key players' previous responses.

How a researcher obtains the interview, how long that interview lasts, the quality and methodology behind the questions asked, whether the interview is on or off the record, whether it is recorded or not, and a host of other factors can deeply influence the quality of the information obtained from interviewees. There are excellent resources that point researchers toward best practices for conducting interviews (Berry 2002; also see this volume). For us, the most important thing a scholar can do is report about the nature of the interviews he or she conducted. Of course, a full description of all these elements for each interview would be time and space consuming. But it helps a great deal if researchers are explicit about key interviewing techniques, such as how the researcher created a record of the interview—through live recording, simultaneous note-taking, or post-event note-taking (and the delay between the interview and the note-taking)—whether interviews were structured, semi-structured, or unstructured; and the length of interviews, especially when the researcher relies on them for critical observations or elements of the analysis.

Absent a video recording (which may be possible in some limited circumstances) or complete transcript of each interview, there is no way for outsiders to assess the quality of the interaction. To that extent, observers must have a degree of trust in the integrity and skills of the interviewer, just as they must trust the capacity of quantitative researchers to accurately code variables and faithfully report results without engaging in data mining. But just as with quantitative projects, observers can be more confident of the process and of the reliability of the results if researchers share as much information about the interview process as possible.

Accuracy of Reporting

Once researchers have conveyed to readers that they have drawn a valid sample and have conducted the interviews in a serious and thoughtful way, they face the task of convincing observers that their reports based on the material reflect the reality of the situation. One simple solution to this dilemma is to post full interview transcripts on a website, so that the curious and the intrepid can verify the data themselves. This ideal standard of qualitative data archiving should be the discipline's goal, and we are not alone in arguing that it should move in this direction (Elman, Kapiziewski, and Vinuela 2010; Moravcsik 2010).

At the same time, however, we fully recognize that it is impractical and even impossible in many cases. It would take significant time and money to transcribe every word of every recorded interview—these are resources most researchers simply do not have. Even if such resources were available, interviews are often granted based on assurances of confidentiality, or are subject to constraints imposed by human subjects research, raising not only practical but also legal and ethical issues (Parry and Mauthner 2004). And, of course, the vast majority of interviews are not recorded at all. These facts inherently constrain the reliability of interview data, but the limitations are not necessarily more severe than analogous problems that plague survey research or data set coding.

Even without providing transcripts, it is possible to communicate the accuracy of reported interview data in a rigorous manner. In many scenarios, the researcher aims to convey that the vast majority of interviewees agree on a particular point. Environmental lobbyists may judge a conservative government unsympathetic to their aims, or actors from across the political spectrum may agree on the importance of civil society groups in contributing to neighborhood policing. Rather than simply reporting this general and vague sentiment, in most instances it is possible to summarize the number of lobbyists expressing this position as a percentage of the total lobbyists interviewed and as a percentage of the lobbyists specifically asked or who spontaneously volunteered their opinion on the government's policy. Similarly, how many policymakers and politicians were interviewed, and what percentage expressed their enthusiasm for civil society groups? This is easiest to convey if the researcher has gone through the process of coding interviews, which is common in some subfields of political science. It is more difficult to present this information if scholars have not systematically coded their interviews; but in these circumstances, it is all the more important to convey a sense of the representativeness of the information cited or quoted.

When quoting an interviewee, researchers are particularly open to the charge that they have simply cherry-picked the most eye-catching statement without regard to its representativeness. To convince readers that such lines are more than a mere adornment to the text, it is vital to communicate whether the quotation represents the average intensity and direction of the response, or whether it distills the most extreme form of reproach or approval. If it is the latter, how many interviewees expressed equivalent intensity in their sentiments? Was that intensity
particularly important as a motivating factor for explaining the actions of key players in the policy debate, even if those players were a numerical minority? In other words, researchers can explicitly convey the meaning of a quotation by addressing its representativeness and the salience of the underlying sentiment for particular actors.

If a researcher's sample frame is sound, if saturation is reached, and if all respondents answer in the same way, we can be relatively certain that the responses have yielded accurate information. But the confluence of these happy circumstances is rare. More often, the sample is incomplete in one or more ways: some people cannot recall key points, others' memories may be inaccurate, or interested parties may misstate facts for private gain. In cases where the interview process is imperfect (in other words, in virtually all real-world cases), it is imperative to report the resulting uncertainty. What percentage of core interviewees were unavailable? How many were unable to answer key questions? What is the range of dissonance among actors over a key point?

These issues are particularly pressing if the goal is to uncover a historical turning point through a process-tracing methodology, such as if a researcher is exploring the genesis of the Iran-contra affair. But they are also central to interview research about less clandestine or event-specific topics, such as if a scholar is focused on the comparative role of doctors in health care reform in advanced democracies. If statements by doctors, policymakers, and other knowledgeable observers suggest disagreement over the relative influence of physicians' associations on policy reform, it is crucial for the researcher to convey this uncertainty.

Once the divergence in opinion is clear, researchers can be explicit about their rationale for trusting some interviews over others and for drawing overall conclusions. There are several common tools and techniques for doing this. In general, evidence drawn from multiple segments of the sample frame is likely to be more reliable than evidence from one part of the sample frame, particularly if a narrow group is arguing for its own significance in the process. For example, if most leaders of doctors' associations assert their groups' importance in the policy process, but if some concur with the vast majority of interviewees from the insurance industry, the pharmaceutical industry, hospitals, and policymakers that physicians' associations were relatively powerless in the process, the latter interpretation is clearly the more credible. In a related manner, information reported against one's interest is almost always more reliable than reports that serve to puff up a particular individual or group. Politicians typically seek credit for good outcomes and try to avoid blame for bad ones. If sitting politicians grant credit for important measures to others—particularly to members of an opposing party—or accept blame when there is nothing to be gained, their statements gain a great deal of credibility.

Beyond reporting the basis for trusting some interviews over others, it is useful to remember that very few studies rely exclusively on interview data for their conclusions. While other types of sources have their own weaknesses, when interview evidence is ambiguous or not dispositive, scholars can fruitfully triangulate with other sources to resolve ambiguities in the record. Perhaps no method of reporting will fully convince skeptics about the accuracy of information gathered through interviews. But strategies such as those suggested here will make evenhanded readers more certain about the reliability of the interview data when judging the rigor of the scholarship and the persuasiveness of the argument.

Forming an Interview Methods Appendix

Keeping in mind these general principles for overcoming some of the most serious challenges to interview reliability, how can researchers quickly and efficiently communicate that they have done their utmost to engage in methodologically rigorous interviewing techniques? We propose the inclusion of an "Interview Methods Appendix" that can be included in any book that relies heavily on interviews. The Interview Methods Appendix can contain a brief discussion of key methodological issues and large amounts of relevant data summarized in an Interview Methods Table. Elements of the Appendix can also be incorporated into articles where space constraints are typically more severe. This is possible in an abbreviated version that follows the main text; or some central elements may be woven into the methodology section of the article itself or—in order to avoid space constraints—as a web link to a full version of the appendix (Moravcsik 2010).

The easiest way to report interview methodology is to break it down into several distinct sections. To illustrate the usefulness of the Interview Methods Appendix and the Interview Methods Table, we use the hypothetical example of a study that explores an attempted ban of the far-right-wing National Democratic Party (NPD) in Germany.

Sample Frame

Report the universe of actors relevant to the study, broken down by subsets to demonstrate a sample frame that draws on a variety of networks and perspectives based on theoretically motivated considerations. In most cases, the author will choose to discuss the construction of the sample frame in the text, especially if interview data are central to her arguments. Even if discussed at length in the text, a short summary of her rationale should appear in the Appendix, perhaps
a paragraph or so long. If the sample frame construction is discussed only in the
Appendix, the discussion should be more extensive.

We believe that explicit reporting of the interview sample frame will improve
readers’ confidence in the results, by displaying the interviewer’s sources openly.
An additional benefit may be through encouraging the interviewer to pay more
attention to sampling strategies.

Example: Interviewees were divided into categories based on their oc-
pupations, with particular attention to political divergence within cat-
egories of politicians, since they were central to the decision to launch
the ban attempt. We sought a diversity of actors from across different
levels of the state—politicians, courts, and bureaucrats—and nonstate
observers of and participants in the process. See the Interview Methods
Table for the breakdown of the categories.

Response Rate and Type
Report the number of interviews sought, obtained, and declined, across each
relevant subset from the original sample frame. This kind of information is so
often lacking, even in scholarship that relies extensively on interviews, yet it
needs to be reported in order for readers to judge the reliability of the interview
data. The author should provide full information about her attempts to gain in-
terview access. This means distinguishing reports of requests that were refused,
merely ignored, or agreed to but then later canceled or refused or simply not
scheduled. It is best to err on the side of providing more information in report-
ing these results. It is also important to note whether the interviews were con-
ducted in person, by phone, by e-mail, or by some other method. See the Inter-
view Methods Table for examples.

Additional and “Snowball” Interviews
Researchers may add interviews without using snowball sampling if they are
unable to get information from respondents from a particular category of the
sample frame. For example, if a researcher were unable to obtain an interview
with her first choice Constitutional Court judge, she may seek an interview with
a second judge as a substitute.

The author should distinguish interviews obtained from “snowball” tech-
niques in two ways. First, the interviewer should report an introduction even if
the “snowballed” interview falls within the sample frame. From our example,
assume that after her first interview with a party politician, a researcher receives
an introduction to a second mainstream party politician that was already in our
researcher’s initial sample frame. The second mainstream party politician
interview should be listed as part of the sample, but the researcher should also
report the connection between the first interviewee and the second interviewee.
That way, readers can assess any potential for bias.

It is also possible, however, that snowball sampling may lead the researcher to
add interviews beyond the originally specified target group, typically because
saturation has not been reached, or because one of the original interviews was
very poor quality, or perhaps because the author realized that a player previously
perceived as minor was in fact central to decision-making and merited an inter-
view. It could also be that the author felt obliged to conduct an additional in-
terview because one of her sample frame interviews insisted she talk to someone
else. No matter the reason, these “snowball” interviews should be reported as such.
The appropriate manner to report this would be in the “Source” column in the
Interview Methods Table (see tables 4.1 and 4.2).

Saturation
Report that saturation has been reached within and across categories, and any
residual uncertainty. Researchers may or may not pursue saturation strategies
depending on the goals of the research. Even if they do not, it is valuable to know
whether saturation has been achieved. If it has, then we are more confident that
additional interviews would not have altered the researcher’s conclusions.

Example: Saturation reached among and across all categories, with the
exception of Far Right party politicians and lawyers, and Constitutional
Court judges.

Format and Length of Interviews
Report the type and length of interviews. Because the broad terms “semi-structured”
or “structured” include wide variation in practice, a footnote at the bottom of
the table should clarify what they mean for the research project at hand. Many
researchers prepare a list of questions and sub-questions to ask in advance of the
interviews. These sometimes are altered or updated throughout the course of the
field research. Unless some other considerations preclude it, the author should
provide a set of the interview questions for “semi-structured” or “structured”
interviews.

Example: Semi-structured interviews in all cases but one, with three
core questions: (1) What were the motives in trying to ban the NPD?
(2) Why did the attempted ban fail? (3) What were the consequences of the failed attempt to ban the NPD? There was also allowance for the interviewees to raise additional issues or comments. The exception was a structured e-mail interview in which the above three questions were posed, with an open-ended fourth question asking for further reflections on the topic.

**Recording Method**

Report whether the interviews were video- or tape-recorded, whether the interviewer took notes during the interview, or whether the interviewer summarized the interview afterward.

Example: Most interviews involved written notes during the interview and were supplemented by extensions on those notes immediately following the end of the interview. In those cases, quotations are the best recollection of the precise phrases used rather than guaranteed verbatim reproductions.

Although for the purposes of clarity and illustration we have summarized some information in points 1 through 6 above, much of this can be easily contained in an interview methods table. In some cases, the interview methods table will have to include additional explication in notes at the bottom of the table or in footnotes. See Table 4.1 for an example of an extended interview methods table.

In addition to the evidence presented in the interview methods table, at least two other kinds of information are needed for observers to gauge the reliability of findings based on interviews.

**Response Rates and Consistency of Reported Opinions or Quotations**

Report the response rates to any key questions or issues flagged in the text of the book or article. Report the representativeness of the opinions or quotations provided.

Example: All respondents answered the three core questions listed under Format and Length. Eighty-five percent of interviewees concurred that the motives for the attempted ban were political, representing at least one person from each subset. Of the 15 percent who thought they
<table>
<thead>
<tr>
<th>INTERVIEWEE</th>
<th>STATUS</th>
<th>SOURCE</th>
<th>SATURATION</th>
<th>FORMAT</th>
<th>LENGTH</th>
<th>RECORDING</th>
<th>TRANSCRIPT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Category 3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constitutional Court judge 1</td>
<td>No response</td>
<td>Sample frame</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constitutional Court judge 2</td>
<td>No response</td>
<td>Substitute in sample frame</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Category 4</strong></td>
<td></td>
<td></td>
<td>Yes</td>
<td>Semi-structured</td>
<td>45 mins</td>
<td>Concurrent notes and supplementary notes w/ 1 hr</td>
<td>Confidentiality required</td>
</tr>
<tr>
<td>Interior Ministry bureaucrat 1</td>
<td>Conducted in person</td>
<td>Sample frame</td>
<td></td>
<td>Semi-structured</td>
<td>45 mins</td>
<td>Concurrent notes &amp; supplementary notes w/ 1 hr</td>
<td>Confidentiality required</td>
</tr>
<tr>
<td>Interior Ministry bureaucrat 2</td>
<td>Conducted in person</td>
<td>Sample frame</td>
<td></td>
<td>Semi-structured</td>
<td>45 mins</td>
<td>Concurrent notes &amp; supplementary notes w/ 1 hr</td>
<td>Confidentiality required</td>
</tr>
<tr>
<td>Justice Ministry bureaucrat</td>
<td>Conducted via e-mail</td>
<td>Referred by Interior</td>
<td>Structured</td>
<td>N/A</td>
<td></td>
<td>E-mail transcript</td>
<td>Confidentiality required</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ministry bureaucrat 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Category 5</strong></td>
<td></td>
<td></td>
<td>Yes</td>
<td>Semi-structured</td>
<td>45 mins</td>
<td>Concurrent notes and supplementary notes w/ 1 hr</td>
<td>Transcript posted</td>
</tr>
<tr>
<td>Constitutional law scholar Schwarz</td>
<td>Conducted by phone</td>
<td>Sample frame</td>
<td></td>
<td>Semi-structured</td>
<td>45 mins</td>
<td>Concurrent notes and supplementary notes w/ 1 hr</td>
<td>Transcript posted</td>
</tr>
<tr>
<td>Constitutional law scholar Berg</td>
<td>Conducted in person</td>
<td>Sample frame</td>
<td></td>
<td>Semi-structured</td>
<td>1 hr</td>
<td>Concurrent notes and supplementary notes w/ 1 hr</td>
<td>Transcript posted</td>
</tr>
<tr>
<td><strong>Category 6</strong></td>
<td></td>
<td></td>
<td>Yes</td>
<td>Semi-structured</td>
<td>1 hr 10 mins</td>
<td>Concurrent notes and supplementary notes w/ 1 hr</td>
<td>Transcript posted</td>
</tr>
<tr>
<td>Far Right scholar Hacker</td>
<td>Conducted in person</td>
<td>Sample frame</td>
<td></td>
<td>Semi-structured</td>
<td>1 hr 10 mins</td>
<td>Concurrent notes and supplementary notes w/ 1 hr</td>
<td>Transcript posted</td>
</tr>
<tr>
<td>Far Right scholar Meyer</td>
<td>Conducted in person</td>
<td>Sample frame</td>
<td></td>
<td>Semi-structured</td>
<td>1 hr 30 mins</td>
<td>Concurrent notes and supplementary notes w/ 1 hr</td>
<td>Transcript posted</td>
</tr>
<tr>
<td>Far Right scholar 3</td>
<td>Not sought</td>
<td>Referred by Far Right</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>scholar Hacker</td>
<td></td>
<td>scholar Hacker</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Category 7</strong></td>
<td></td>
<td></td>
<td>Partial</td>
<td>Semi-structured</td>
<td>55 mins</td>
<td>Concurrent notes and supplementary notes w/ 1 hr</td>
<td>Confidentiality requested</td>
</tr>
<tr>
<td>Expert journalist 1</td>
<td>Conducted by phone</td>
<td>Sample frame</td>
<td></td>
<td>Semi-structured</td>
<td>55 mins</td>
<td>Concurrent notes and supplementary notes w/ 1 hr</td>
<td>Confidentiality requested</td>
</tr>
<tr>
<td>Expert journalist 2</td>
<td>No response</td>
<td>Sample frame</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Category 8</strong></td>
<td></td>
<td></td>
<td>Yes</td>
<td>Semi-structured</td>
<td>1 hr 10 mins</td>
<td>Audio recording</td>
<td>Redacted transcript posted</td>
</tr>
<tr>
<td>Antifascist NGO leader Korn</td>
<td>Conducted in person</td>
<td>Sample frame</td>
<td></td>
<td>Semi-structured</td>
<td>1 hr 10 mins</td>
<td>Audio recording</td>
<td>Redacted transcript posted</td>
</tr>
<tr>
<td>Antifascist NGO leader Knoblauch</td>
<td>Conducted in person</td>
<td>Sample frame</td>
<td></td>
<td>Semi-structured</td>
<td>30 mins</td>
<td>Audio recording</td>
<td>Redacted transcript posted</td>
</tr>
<tr>
<td>Antifascist NGO leader 3</td>
<td>Not sought</td>
<td>Referred by Antifascist</td>
<td></td>
<td>Semi-structured</td>
<td>30 mins</td>
<td>Audio recording</td>
<td>Redacted transcript posted</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NGO leader Korn</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anti-discrimination NGO leader Spitz</td>
<td>Conducted in person</td>
<td>Referred by Antifascist</td>
<td>Semi-structured</td>
<td>1 hr 30 mins</td>
<td>Audio recording</td>
<td>Transcript posted</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>NGO leader Korn and</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>by Far Right scholar</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Meyer</td>
<td></td>
<td>Meyer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>High</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>See <a href="http://www.bleichpekkanen">www.bleichpekkanen</a> .transcripts</td>
</tr>
</tbody>
</table>
were problem-driven (i.e., a response to the objective rise of racism in German society), all came from the NGO sector, with one coming from the scholarly sector. The quotation from the SPD politician in the text is representative of the average response to this question among the 85 percent of interviewees who viewed the motives as political.

Confidence Levels and Compensation Strategies

Report concerns about uncertainty over particular points, explicit attempts made to reduce uncertainty, and any residual uncertainty.

Example: Absent interviews with Far Right politicians, we devoted a portion of our interviews with scholars of Far Right political parties and with the expert journalist to ask their opinions of Far Right party perspectives on the attempted ban. Their direct contact with Far Right leaders allowed them to speak with moderate confidence about those views, increasing our certainty about the positions of the NPD in particular. Similarly, lacking interviews with Constitutional Court judges, we tracked down scholarly articles and previous court opinions about Far Right parties by key judges. In both these cases, our certainty about their perspectives cannot be complete, but we are confident that the missing information does not inordinately bias our findings given the convergence of information and saturation obtained from other sources.

The discussion we have provided above applies most directly to researchers who use interviews as an integral part of their information gathering. Yet most scholars use interviews as only one source of information among many, and some use interviews quite sparingly as a way to supplement information gathered from written material. In these cases, the interview methods appendix and interview methods table can easily be adapted. For example, it may be that a researcher conducting a similar study had access to previous work that included internal political party documents revealing strategies and approaches to the attempted ban of the NPD, credible investigative journalism that uncovered the attitudes of Constitutional Court judges, book-length assessments of the topic from expert journalists, constitutional scholars, and Far Right, scholars and multiple policy statements of NGOs that participated in the process.

Supplementing this extensive record with additional interviews from categories 1–3 and 5–8 may be extremely useful. But it may not be necessary. Here, the scholar can discuss her sample frame and note that saturation has been achieved.
from written sources across all but one category. The interview methods table would then serve to demonstrate the type, character, and extensiveness of the limited set of interviews that were conducted, as well as to highlight the rigor of the scholar's approach to using interviews as part of a well-planned research strategy. In this case, the table would look like Table 4.2.

We do not argue that interviews are the only or the best source of information. However, in line with the contents of this volume, we think there are cases when interview data can contribute greatly—and perhaps uniquely—to research. This motivates our attempt to bolster the trustworthiness of interviews as a valued source of data. We certainly expect that most scholars will use other streams of evidence besides interviews—be it secondary sources or data analysis—to construct their arguments. In such mixed-methods approaches, the researcher can and should signal when sources besides interviews are used to gain information needed to round out the sample frame.

We recognize that legitimate concerns may force researchers to keep some details of the interview confidential and anonymous. We all must respect constraints imposed by institutional review boards, by informants themselves, or by professional ethics. In certain cases, the interview methods appendix may contain "confidentiality requested" and "confidentiality required" for every single interview. We do not seek to change prevailing practices that serve to protect interview subjects. However, we believe that even in such circumstances, the interviewer can safely report many elements in an interview methods appendix and interview methods table—to the benefit of researcher and reader alike. Whether or not scholars are at liberty to provide video recordings of every interview in an online "appendix" or through a qualitative data archive, they should still report elements such as their sample frame, nonresponse rates, format of interviews, use of snowball sampling, and confidence levels. A consistent set of expectations for reporting these will give readers more confidence in research based on interview data, which in turn will liberate researchers to employ this methodology more often and with greater rigor.

To the extent that researchers are able to provide transcripts of their interviews in online appendixes or qualitative data archives—perhaps following an initial embargo period standard among quantitative researchers for newly developed data sets, or for archivists protecting sensitive personal information—there are potentially exponential gains to be made to the research community as a whole (Elman, Kapiszewski, and Vinuela 2010). Not only will this practice assure readers that information sought, obtained, and reported accurately conveys the reality of the political or policymaking process in question; it will also allow researchers in years to come access to essential interviews with key practitioners that would otherwise be lost to history. Imagine if in forty years a scholar could reexamine a pressing question not only in light of the written historical record, but also with your unique interview transcripts at hand. Carefully documenting interviewing processes and evidence will enhance our confidence that we truly understand political events in the present day and for decades to come.
Contents

Preface vii
List of Contributors ix

Introduction. “Just Talk to People”? Interviews in Contemporary Political Science
Layna Mosley 1

Part 1 GENERAL CONSIDERATIONS: RESEARCH DESIGN, ETHICS, AND THE ROLE OF THE RESEARCHER 31
1. Aligning Sampling Strategies with Analytic Goals Junia F. Lynch 31
2. The Ethical Treatment of Human Subjects and the Institutional Review Board Process Sarah M. Brooks 45
3. The Power of the Interviewer Lauren M. Macklen 67
4. How to Report Interview Data Erik Bleich and Robert Pekkanen 84

Part 2 ADDRESSING THE CHALLENGES OF INTERVIEW RESEARCH 109
5. Crafting Interviews to Capture Cause and Effect Cathie Jo Martin 109
6. Using Proxy Interviewing to Address Sensitive Topics Melani Cammett 125
7. Working with Interpreters Lee Ann Fujii 144
8. The Problem of Extraterritorial Legality William Reno 159