Why Do We Speak to Experts? Reviving the Strength of the Expert Interview Method

Christian von Soest

In political scientists' drive to examine causal mechanisms, qualitative expert interviews have an important role to play. This is particularly true for the analysis of complex decision-making processes, where there is a dearth of data, and for linking macro and micro levels of analysis. The paper offers suggestions for making the most effective and reflective use of qualitative expert interviews. It advocates an encompassing, knowledge-based understanding of experts and argues for the incorporation of both "inside" and "outside" experts, meaning those that make and those that analyze political decisions, into an integrated analytical framework. It puts forward concrete advice addressing this technique's inherent challenges of selecting experts, experts' personal biases, and the systematic capturing of evidence. Finally, the article suggests that the combination of expert interviews, experimental methods, and online interviewing can meaningfully strengthen the evidentiary value of this important data collection technique.

Why Do We use Expert Interviews?

Experts are the observers of and mechanics behind what social scientists call "causal mechanisms" (Elster 1998). Where researchers often gather data of onset points and examine the outcomes of political decisions, rare are the possibilities of looking into the inner workings of political processes. Here interviewing experts becomes key to understanding how "X and Y interact" (Gerring 2017, 45). In this paper I provide suggestions on how the important data collection technique of qualitative expert interviews can be used in the most effective manner to discern decision-making and institutional behavior.

Current "quantitative and causal inference revolutions"[1] (Pepinsky 2019, 187; for a critique, see Elman and Burton 2016) have dominated the methodological debate in the social sciences. Scholars conduct causal investigations with experimental or quasi-experimental designs (Dunning 2012) and with observational time-series cross-sectional data (Blackwell and Glynn 2018).[2] These developments have led to a shift in mainstream political science research toward investigating micro-level processes and from questions of external validity to ones of internal validity (Pepinsky 2019). This application of rigorous methods of causal inference has greatly expanded our knowledge about political behavior and its outcomes. However, despite the fact that quantitative approaches have driven the recent methodological debate in comparative politics and international relations,[3] the data collection technique of qualitative expert interviews still has an important and arguably even growing role to play "within a discipline addicted to causation" (Anderl 2015, 2).

As a tool to investigate causal mechanisms, expert interviews hold three advantages. First, they may add to experimental findings about micro processes and how decisions were made in practice (Fu and Simmons 2021). Researchers can blend general results with context-specific information that is often not in the public domain. In this way, qualitative information from expert interviews facilitates the interpretation of correlational analysis and can thereby improve causal inference by statistical means (Gerring 2017, 44–45; Glynn and Ichino 2015; Kabeer 2019). Second, political science recurrently deals with "big" questions that do not lend themselves

Christian von Soest (ID) *is a Lead Research Fellow and Head of the Peace and Security Research Program at the German Institute for Global and Area Studies (GIGA) in Hamburg, Germany (christian.vonsoest@giga-hamburg.de). His research focuses on international sanctions and interventions, authoritarian regimes, legitimation strategies, qualitative methods, and African politics. He has published in leading publication outlets such as* Democratization, European Journal of Political Research, European Political Science Review, International Studies Quarterly, *the* Journal of Modern African Studies, *and* Journal of Peace Research.

easily to experimental or statistical analysis, particularly in cases where the number of observations is low. Experts can aggregate and weigh different pieces of information (refer to the section "Using expert interviews"). Third and relatedly, expert interviews can provide the data to link the macro and micro levels of analysis. For instance, in peace and conflict research interviewing experts may shed light on the impact of insurgents' international sponsors on local peacebuilding dynamics (Balcells and Justino 2014) or on how the mobilization of individuals in civil war is linked to social structures (Shesterinina 2016).

However, despite its high practical relevance and the existence of classic works on the qualitative interview method (Aberbach and Rockman 2002; Dexter 2006 [1970]; Tansey 2007), current reflection on the promises and limits of this technique is sparse. There is also little guidance offered on how to use the method most effectively (Fujii 2017; Lareau 2021). It is my key contention that pursuing an encompassing understanding of "experts," one based on the knowledge of individuals, allows for the construction of a systematic "architecture" (Trachtenberg 2006, 32) of data sources and increases expert interviews' analytical value.

Oftentimes, researchers use qualitative expert interviews as a key method to gather information about political processes but only mention the technique in passing. The resulting problems are pertinent: 1) the selection of experts regularly does not follow clear guidelines; 2) experts' personal biases are often not tackled; and 3) evidence is not captured systematically. These omissions severely inhibit the potential of qualitative expert interviews to trace causal mechanisms, which is their key potential strength.

In response, I seek to contribute to systematically applying qualitative expert interviews in a more effective manner. I situate these recommendations in respect to two particular bodies of work: a) an own research project on political interventions into the tax administration; and b) examples from the literature on political violence and civil wars. The paper's contribution is threefold: My first argument is that for most theoretical and empirical research questions of interest, social scientists should focus on a broad notion of experts. When choosing experts, researchers should select both "inside" and "outside experts" and include them in an integrated framework. Second, I propose concrete measures for dealing with the technique's main challenges of selection of experts, personal biases, and systematic capturing of evidence. Third, I argue that the combination of expert interviews with other methods such as list experiments and online interviewing are key means to mitigate practical challenges and to reduce social-desirability bias. These strategies serve to make qualitative expert interviews even more effective research tools.

Using Expert Interviews

Broadly understood, experts have specific knowledge about an issue, development, or event. Hence, following Dexter's (2006) classic understanding, an expert is any person who has specialized information on or who has been involved in the political or social process of interest. Consequently, I advocate an encompassing notion of experts. They "might be academics, practitioners, political elite, managers, or any other individuals with specialized experience or knowledge" (Maestas 2018, 585). Tapping into their insights responds to the fundamental challenge that a lot of issues of interest in political science and in the social sciences more generally are not directly observable, documented, or made transparent.

Most publications on qualitative interviews in political science focus on elites rather than experts. However, an encompassing notion of experts allows us to explicitly situate their status, knowledge, own interests, and potential biases in relation to each other, and to construct a coherent "architecture" of interview sources. With this "realist approach," expertise is based on real knowledge (Collins and Evans 2007, 3). Following Hafner-Burton et al. (2013, 369; also Pakulski 2008), elites "occupy top positions in social and political structures" and "exercise significant influence over social and political change." This status-oriented perspective is sensible for research that focuses on the perspective of a particular set of decision makers, for instance members of parliament (on Russia, see Rivera, Kozyreva, and Sarovskii 2002). However, members of the elite are not necessarily knowledgeable about the political process or event of interest. In this case, they do not qualify as experts and the researcher should desist from selecting them as interview partners. For most research objectives, the selection of expert interview partners should therefore be problem- and expertise-centered rather than status-oriented.[4] Yet selection based on expertise and on status may overlap. In other words, individuals can both hold expertise and be members of an elite group.

We can distinguish three broad functions of expert interviews: to 1) (inductively) explore a research topic/generate hypotheses; 2) collect data for qualitative and mixed-method designs (for instance to situate findings from experimental studies in real-world environments); and 3) generate quantitative data and allow for statistical inference. Based on the existing literature, I differentiate four main applications of qualitative expert interviews: assessment, aggregation, anticipation, and affirmation.[5] The first and most relevant is *assessment*, meaning that experts share their judgement on political and social processes. Most importantly, assessment entails that experts in an analytical way "reconstruct an event or set of events" (Tansey 2007, 766), often providing the empirical basis for the process tracing method (Beach and Pedersen 2019; Bennett and Checkel 2015).[6] Thus, the unit of analysis for expert interviews is a particular event, development, or decision making process.

The second, related function is *aggregation*, as experts are well-suited to reducing real-world complexity and

bundling together multifaceted phenomena. Compared to the assessment function, aggregation can also entail the reconstruction of events and provision of information, but in a more descriptive manner. Third, experts can use their research or personal experience for *anticipation* and the prediction of events, of actors' behavior, and/or of long-term developments—for instance the propensity for future violent conflict (Hanson et al. 2011; Meyer, De Franco, and Otto 2019). However, the accuracy of expert prediction vis-à-vis other methods, in particular data-based approaches, is highly contested (Hanson et al. 2011; Hegre et al. 2021; Tetlock 2005).

Finally, expert interviews may serve as a method of *affirmation*, meaning the confirmation or disproving of prior research results, information from other sources, or anecdotal evidence (Tansey 2007, 766–67). This function has a dark side to it: selectively invoking expert insights to support a partisan purpose or to undergird political action potentially distorts results and is unethical. As I will outline with examples from a research project on particularistic interventions into the tax administration as well as from the literature on political violence, purposefully selecting experts, dealing with their potential biases, and systematically relating different expert interviews to each other are key means of avoiding this pitfall.

The Selection of Experts

Qualitative expert interviews clearly lend themselves to purposeful, non-probability sampling (Goldstein 2002; Tansey 2007), as expert judgements are inherently personal and not necessarily representative or replicable. Prior thorough reflection on the knowledge, but also potential information gaps and personal biases of experts must therefore guide the interviewee selection process. For this, I suggest a) to integrate inside and outside experts into one common analytical framework and b) to not only focus on high-level but also lower-level inside experts.

Inside versus outside experts: Most fundamentally, I suggest selecting both experts who make decisions ("inside experts") and ones who analyze them ("outside experts") as respondents (see table 1). Inside experts are decision makers who actually shaped the political or social process of interest. In this case, expert interviews are a proxy for participant observation (Pouliot 2015, 247). Former tax administration officers, rebels, members of Congress, or international organization representatives can generally also be considered "insiders." In fact, they oftentimes represent particularly reliable sources as they are more inclined to provide information after they have left the organization in question.[7] In contrast, external experts are outsiders to the process in question. They gain their expertise through research, experience, or interaction with policymakers and officials who took a decision. In contrast to insider experts, they themselves are not the object of analysis. Table 1 introduces guidelines for distinguishing inside from outside experts.

The categorization of inside and outside experts always needs to be made with respect to the particular research agenda and the function of expert interviews for the respective project. If political interventions into the tax administration are the focus of research, then respondents from the ministry of finance as well as tax officials on different hierarchical levels would clearly be regarded as inside experts, while taxpayers, advisors, and social scientists would be considered outside experts.

The differentiation between insiders and outsiders has key implications for expert interviews' analytical value and the interpretation of the insights provided (Beach and Pedersen 2019, 207–9). Interviewing inside and outside experts holds promises but also pitfalls (see the summary in table 1): First of all, while inside experts may provide detailed first-hand accounts and "hidden" knowledge that is not publicly available, they also have a stronger interest in withholding or molding information to "look good"

Table 1
Inside and outside experts

	Inside Expert (Actor)	Outside Expert (Analyst)
Object of analysis/directly involved in the process in question	Yes	No
Status/occupation	Policymakers, administrators, (but also) mid-/low-level civil servants, activists, citizens	Researchers, policy analysts
Main use	Assessment, affirmation	Aggregation, assessment, anticipation, affirmation
Advantages	Primary source, direct participants'/ decision makers' perspective	Secondary source, higher potential for prior reflection, "big picture"
Disadvantages	Memory lapses, distortion of information, systematic bias	Distance to event/process of interest, perspective may be based on distorted sources

and to inflate or downplay their influence. In contrast, outside experts may be in a better position to provide the "big picture," meaning the *assessment* of processes and events as well as the *aggregation* of various sources. On the downside, they themselves rely on information from others and might lack knowledge of how exactly choices were made (e.g., which emotions were involved in a particular decision).

Given these advantages and disadvantages, I suggest that ideally researchers select *both* inside and outside experts and include them in an integrated analytical framework. The heuristic differentiation of both groups allows us to relate them to each other systematically. For their purposeful selection through the recommendations of other experts, snowball sampling and related techniques can be used (Goldstein 2002; Heckathorn and Cameron 2017; Shesterinina 2016, 415). Outside experts should be selected on the basis of their publication records, their (local) expertise, and, if applicable, their prior work for high-quality country-based indices such as the Bertelsmann Transformation Index (BTI 2021).

The case for mid- and low-level insiders: Oftentimes, scholars disregard the influence and informational advantage of inside experts who do not occupy top positions—be it election campaigners, civil society activists, or mid-level civil servants. For example, to investigate political intervention into the tax administration in two African countries (von Soest 2009, 47), I interviewed tax officers from the operational level and systematically related their perspectives to those from the tax agency leadership. It was thereby possible to contrast the information provided by the two groups and to get evidence on both the interference in the hiring and career trajectories as well as the day-to-day work of those officers who actually collect taxes and conduct tax audits on the ground. Respondents mentioned that several companies owned by government politicians had never been audited and at no time paid tax. "A common phrase employees heard from their managers in these cases was 'don't go to this man, this man is difficult'" (von Soest 2009, 107). On the other hand, the interviews pointed to instances of resistance and the insistence on preserving administrative standards.

As George and Bennett (2005, 103) state: "Often, lower-level officials who worked on an issue every day have stronger recollections of how it was decided than the top officials who actually made the decision but who focused on the issues in question only intermittently." In addition, lower-placed individuals are often less trained to provide a "polished" version of events. They are thus regularly more helpful in reconstructing political and social processes at the micro level. Thus, if possible, researchers should strive to interview inside experts—ones involved in the process in question—on different hierarchy levels (primary sources), in addition to outside experts (secondary sources).

In addition to the selection of both inside and outside experts, it is of crucial importance to limit the possibility that respondents' personal biases compromise the results and thereby limit expert interviews' analytical value.

Dealing with Personal Biases

Generally, a semi-structured format with defined topics and preformulated questions is most appropriate for expert interviews (Tansey 2007). They are guided by clear themes, keywords, and established questions while simultaneously allowing for follow-up enquiries and probes. In this way, they represent a useful combination of structure and flexibility, and facilitate both comparability and context sensitivity. To realize the full potential of this powerful data collection tool, I now put forward suggestions on how scholars can actively tackle two core personal biases when conducting expert interviews: reactivity and subjectivity.

Reactivity: The Relational Nature of Expert Interviews

Social scientists have long discussed the fact that individuals might adjust their behavior and utterances when participating in studies (Orne 1962). Scholars—particularly those working with an interpretivist orientation—have stressed the "relational nature" of the interview process, meaning that both expert interviewees and researchers influence its outcomes (Berry 2002; Dexter 2006; Fujii 2017; Peabody et al. 1990). An extremely insightful literature has reflected on the positionality of academics (often from the Global North) and of local interviewees, particularly in post-conflict contexts (e.g., Krause 2021; Mwambari 2019).[8] By design, the expert interview is a personal, face-to-face conversation between at least two individuals. It is principally asymmetric in that the researcher poses questions and the expert provides their assessment.

Yet interviewing needs to be understood as a two-way dialogue rather than a one-way interrogation (e.g., Cramer 2016; Fujii 2017). Doing research on the tax administration's political environment in two African countries, even seemingly neutral outside experts (social scientists and employees of international organizations) were understandably well aware of their answers' implications: a) despite prior information, they repeatedly asked me about the background to my research ("Are donors funding your project?"); b) some experts inquired about my and others' judgements ("You have certainly done a lot of other interviews—what did they say?") (von Soest 2009).

The interviewer influences the nature of the interview in at least two ways: by who they are and by how they pose questions. First, the researcher's status, years of experience, gender, age, nationality, and further personal traits may affect the answers respondents give. Experienced researchers have an advantage in establishing their expertise—such as on the organization of the tax administration,

United Nations peacekeeping missions, or the structuring of militant rebel groups—and consequently are taken seriously by the respondent, while inside experts—be they tax officers, soldiers, or (former) insurgents—might conversely see young academics as "harmless" outsiders and therewith more readily share information with them (Autesserre 2014, 284–86). Researchers should therefore reflect on these matters ex ante and transparently report on how their positionality vis-à-vis experts—particularly insiders—might have influenced the interview situation and the information provided (George and Bennett 2005, 99).

Second, the questions posed and the interviewer's own reactions may affect respondents' answers. It is generally accepted that to elicit informed and unbiased insights, scholars should take a neutral stance and work with nonsuggestive questions. Yet the reactive nature of expert interviews works to the researcher's advantage in a twofold manner. First, they can work with probes—following up on answers or making reference to other interviews or sources—to address contradictions. Furthermore, the researchers can use various strategies to obtain all the information the respondent is willing to share (Dunning 2015, 233). Using an approach from criminal investigation, researchers can for instance pose similar follow-up questions using different wordings and formulations. This strategy proves particularly useful when tracing sensitive issues such as political intervention into the tax administration or war crimes (von Soest 2007; Fujii 2010). Second, the researcher can use the respondent's posture and (nonverbal) reactions to questions like nodding, hesitation, gestures, or expressively strong affirmation as a further source of information. This interview "meta-data" (Fujii 2010, 231) helps to trace the meaning respondents attach to events or processes (Pouliot 2015).

Subjectivity: Lacking Memory and Active Misrepresentation

Even more so than other data sources such as archival material, expert interviews are acts of "purposeful communication" (George and Bennett 2005, 99). The information provided is always subjective and colored by the experts' worldviews, interests, employment status, and cognitive abilities. Furthermore, as Berry (2002, 680) soberingly notes, "it is not the obligation of a subject to be objective and tell us the truth." Experts can intentionally or unintentionally mispresent information (see table 2 for different potential biases).

This problem is particularly acute for internal experts, and even more so for those who are highly exposed to public scrutiny such as policymakers (Trachtenberg 2006, 154). Yet external experts also have their ideological and personal predispositions that affect the answers they give. Analyzing highly contentious issues such as the support for

Table 2
Potential biases

1) Purposeful misrepresentation
2) Unintentional misrepresentation
3) Subjective interpretation
4) Lacking or insufficient memory/knowledge

violence or working in post-conflict and authoritarian environments makes social-desirability bias particularly salient (Lyall 2015, 204; Tripp 2018). Tax officers hardly instantly or directly spoke about corrupt practices in the tax administration (von Soest 2007). It also brings security concerns (for both respondents and researchers) and fundamental ethical issues to the fore (Clark 2006; Fu and Simmons 2021; Parkinson 2022; Wood 2006). Furthermore, timing plays a crucial role; obviously, the further researchers go back in history, the higher the probability that respondents cannot remember certain events or processes.

Finally, respondents differ in their interpretations. The "Rashomon effect" denotes the simple fact that different participants in a process have alternate views as to what actually took place (George and Bennett 2005, 99f.). In consequence, the evidentiary value of just one expert interview is close to zero; it needs to be cross-checked with other interviews or data streams ("triangulation," discussed later). I would furthermore advise scholars to conduct an assessment of the motives a respondent might have to distort or conceal information before interviews start (Beach and Pedersen 2019, 210).

Combination with Other Data Collection Techniques/ List Experiments

To overcome or at least limit the inherent dangers of personal bias, scholars can gain from combining qualitative expert interviews with other data collection techniques. My focus here lies on highly structured experimental methods; in my view, existing research has hardly explored the potential to integrate them with expert interviews. Experimental methods reduce incentives to distort or withhold information, particularly when the interviewer investigates controversial topics. For instance, while survey-based conflict research that directly asks about respondents' *exposure* to violence is increasingly common, posing upfront questions on *attitudes to violence* is not (Balcells and Justino 2014). Social-desirability bias regularly renders responses invalid. To counter this, scholars have designed so-called list experiments to help examine "true biases and preferences that would be otherwise difficult to reveal" (Dietrich, Hardt, and Swedlund 2021, 603).

In a list experiment, the researcher asks the respondents how many preformulated statements they consider correct. These statements are taken from a list of assertions

that includes a sensitive item, such as support for a violent organization (Swedlund 2017). One group of interviewees receives the complete list of statements, while the other group gets the same one but without the sensitive statement. The comparison of the two groups' mean answers provides "an estimate of how many respondents believe the sensitive item to be true" (Swedlund 2017, 471).[9] The precondition for applying this indirect assessment method is a sufficiently high number of interviewees (Lyall 2015). Yet even if the number of respondents is too low to meaningfully calculate the difference between the two groups' mean responses, simply comparing them gives an indication about the interviewed experts' true preferences and the level of misrepresentation at play. Standardized methods such as these can be easily integrated into interviews with inside experts, particularly on contentious issues. For instance, in her research on the perceptions of officers at NATO headquarters, Hardt (2018) started each interview with a five-minute paper-and-pen survey experiment. In this way, the cross-fertilization with standardized data collection techniques can strengthen the evidentiary value of expert interviews, in particular those conducted with policymakers (inside experts).

Capturing Evidence

In addition to systematically selecting experts as well as dealing with the reactivity and subjectivity of the expert-interview method, the capturing of evidence should follow clear guidelines. Researchers should ensure that the material used is representative of the whole empirical corpus and triangulate expert interviews.

Representativeness: The Power of the Good Quote

Regularly, scholars use extracts or snippets from interviews to capture a certain aspect of the empirical reality in a concise and vivid manner. In doing so, researchers should relate these extracts to the whole empirical corpus and thereby counter the often-made accusation that quotes from expert interviews are "cherry-picked" (Dunning 2015, 232; Elman, Gerring, and Mahoney 2016, 383; Tripp 2018, 731, 735). This is all the more important as due to ethical and methodological constraints qualitative research is much harder to replicate than statistical analysis. Often, expert-interview transcripts cannot be made public.[10]

To overcome the power of the good quote and avoid biases in using references, scholars should clearly catalogue the procedures guiding the aggregation and interpretation of information (Tripp 2018, 735–36). Two aspects are of crucial importance here. First, as outlined earlier, it starts with the balanced selection of inside and outside experts. As Schedler (2012, 31) notes, the "quality of expert judgments … depends, first of all, on the quality of experts." Second, when using extracts researchers should always state how strong expert consensus and how

representative a particular quote is compared to all interview statements. Sentences such as "none of the outside experts interviewed for this study maintained that" or "this sentiment was widely shared among civil servants" (von Soest 2009) clearly situate the quote in the empirical corpus of expert interviews and other data sources.

With their comparatively high level of structuring, semi-structured expert interviews permit us to systematically assess the degree of "inter-expert agreement" (Dorussen, Lenz, and Blavoukos 2005, 325), therewith putting the representativeness of selected interview extracts on firm grounds (Dunning 2015, 215, 232). The level of inter-expert agreement can then be reported as a measure for the reliability and validity of the obtained information. In doing so, the threshold for inter-expert agreement is considerably higher for outside experts than for inside ones.

As a general principle, the number of expert interviews should be as high as possible (Dorussen, Lenz, and Blavoukos 2005; on expert surveys, see Maestas 2018). A large corpus of expert interviews facilitates situating individual perspectives, and thereby better gauging the representativeness of individual pieces of information. Of course, this requirement is dependent on: a) the relationship to other data streams (is information from other sources available?); b) the nature of the topic/research interest in question (for instance, hypothesis generation versus hypothesis testing); and c) whether the researcher deals with "hard-to-survey populations" (Tourangeau 2014). Hence, this guiding principle should not make scholars reject the method in difficult cases.

External and Internal Triangulation

The classic strategy to enhance confidence in the accuracy of the information from expert interviews is to blend it with other data streams (Denzin 1978). Expert interviews are rarely used as a stand-alone technique, but form part of an architecture of sources. My comparative study on political interventions into the tax administration in two African countries relied on over 150 semi-structured expert interviews that were complemented with the annual tax-administration reports, data on revenue performance, and secondary literature to strengthen the process tracing analysis (von Soest 2009; for another example, Lundgren 2020). The precondition for this triangulation is that the sources do not depend on each other (Beach and Pedersen 2019, 215).

In addition to the established triangulation of different data sources, I argue that more can be made of "internal triangulation"—meaning the organized cross-checking of information collected via expert interviews themselves. Analysis should systematically consider different expert groups to achieve as much control as possible. This includes both inside (actors) and outside (analysts) experts

as well as, if applicable, insiders on different hierarchical levels. This strategy is key to adequately validating the information provided. To investigate political intervention into the tax administration (von Soest 2009, 47), internal control was sought by interviewing tax officers from the leadership versus from the operational level (middle- and even low-level tax officers). Indeed, the leadership overall painted a rosy picture, while tax officers in one country reported concrete interventions that politicians targeted their day-to-day work with. In addition, experts from outside the tax office were interviewed. These outsiders were businesspersons, tax advisors, civil society representatives, policymakers, and social scientists. The five respondent groups allowed me to control for the internal perspective and provided additional insights (von Soest 2009, 48).

This categorization of experts furthermore helps to refer to respondents while preserving their anonymity. As a convention, researchers could systematically designate sources as "insider" or "outsider" in using quotes and analyzing data. For instance, "According to inside expert A" or "Representing the majority of assessments, outside expert E stated that". This categorization may complement or even supplement specific positional descriptions (which might at times be too revealing), such as "a mid-level tax officer" or "a leading activist."

Online Expert Interviews

Expert interviews are designed to be in-person encounters. The face-to-face interview situation allows one to establish rapport with the respondent, notice cues, and record further non-verbal meta-data. Researchers are able to quickly adapt to the interview situation and flexibly pose follow-up questions. This ensures a high degree of validity. Also, being "in the field" eases or even is a precondition for the purposeful, encompassing, and balanced selection of respondents through referrals and snowball sampling (e.g., Driscoll 2021; Goldstein 2002; Heckathorn and Cameron 2017). Yet in authoritarian and post-conflict environments access to the field is fundamentally restricted (de Vries and Glawion 2021; Wackenhut 2018). More generally, conducting interviews on controversial topics might create serious risks for both respondents and researchers, even in otherwise peaceful contexts (Irgil et al. 2021). Most recently, the COVID-19 pandemic has drastically inhibited travel as well as access to individuals, whether within one's country or abroad (Mwambari, Purdeková, and Bisoka 2021; Schirmer 2021).

Yet digital research methods have emerged as a viable alternative to in-person meetings. In particular, synchronous online interviewing through video-conferencing tools like Zoom or MS Teams are a useful supplement or even (if necessary due to security or other risks) replacement for in-person expert interviews. Scholars have identified particular methodological, security-related, and

ethical challenges regarding remote meetings, such as the increased probability of sampling bias (some experts can hardly be reached through the Internet), security agencies' online surveillance, and complicated trust-building in the absence of personal interaction. This could make expert interviews more superficial (Irgil et al. 2021, 1513; Mwambari, Purdeková, and Bisoka 2021; van Baalen 2018). Selecting interviewees and creating trust should be particular problems for decision makers; less so for external experts who routinely analyze the processes and events in question and are used to digital communication.

Recent experiences provide greater optimism and indicate that these challenges can be overcome. First, scrutiny of both researcher and participant perspectives suggests that scholars can also establish rapport via synchronous video-conferencing tools that transmit audio and pictures. Both interviewers and interviewees may feel comfortable using such technology (Archibald et al. 2019; Lo Iacono, Symonds, and Brown 2016). Even referrals to further respondents seem possible; however, anecdotal evidence suggests that a prior on-site stay eases the selection of experts significantly (Schirmer 2021). In addition, online meetings are less costly than in-person research (e.g., Irgil et al. 2021, 1513). The number of expert interviews can thereby be increased, and online conversations also be used as a follow-up to prior in-person meetings.

Thus, using digital technology for conducting expert interviews can supplement traditional in-person meetings. As for in-person interviews, while more difficult to achieve, it is imperative that researchers select a balanced sample of outside and, where available, inside experts. They should systematically reflect on and report potential biases in the information and assessments gained through online data collection techniques, and how these might have influenced findings.

Conclusions

Much of the current mainstream methodological discussion in political science has centered around the use of experimental methods and new ways of making causal inferences via statistical means. Further insightful research has discussed the methodological, psychological, and ethical issues of field-research methods such as immersion (most recently, Driscoll 2021; Irgil et al. 2021; Krause 2021; Parkinson 2022). Yet considerations specifically focused on the data collection technique of qualitative expert interviews have attracted less attention.

In this paper I have made concrete suggestions on how to conduct expert interviews in a structured and transparent manner to strengthen their validity and reliability. This is all the more important as qualitative evidence—such as that from expert interviews—is much harder to replicate than statistical data. I argued that inside (actors) and outside (analysts) experts should both be included in an integrated analytical framework. To deal with expert

interviews' inherent challenges of selecting experts, countering their potential biases, and systematically capturing evidence, it is imperative to: a) reflect on the relational nature of the interview situation and to disclose one's positionality; b) triangulate interviews and compare them with other data streams (internal and external triangulation); and c) systematically situate quotes in relation to the whole empirical corpus. Methods such as list experiments can reduce social-desirability bias, particularly for interviews with inside experts, while interviewing online is a useful supplement to traditional in-person meetings.

Conducted in such a structured manner, qualitative expert interviews have a crucial role to play not only in the current drive to assess causal mechanisms but also in generating important descriptive insights about the "what" and "how" of political processes and events (Fu and Simmons 2021). They help to examine how actors actually behaved in a decisional context and to discern what happened and why it happened. Findings from expert interviews thereby contribute to analyzing political science's "big" questions and to linking the macro and micro levels of analysis. Future publications can present more in-depth advice on the structured analysis of expert-interview data and also formulate suggestions on how to bring journals' preregistration and transparency requirements into line with this qualitative method's specificities. Further discussing their systematic application will be crucial to making the most of expert interviews as a key data collection tool for social scientists.

Acknowledgement

Notes

1 Angrist and Pischke (2010) also dub this the "credibility revolution."
2 These design-based causal-inference models largely stem from behavioral economics (Angrist and Pischke 2010) and development economics (Banerjee and Duflo 2009).
3 In addition, in what can be called "political ethnography" (Schatz 2009), there is a vivid and extremely insightful debate about the methodological aspects, practicalities, and ethics of field research (Irgil et al. 2021; Krause 2021; Parkinson 2022; Wedeen 2010; see also, Driscoll 2021).
4 Generally speaking experts are selected purposefully, while on the other hand a randomly selected individual would not be considered an expert.
5 This draws on Tansey's differentiation (2007, 766–67), who introduced four uses of *elite* interviews: "corroborate other sources"; "establish what a set of people think"; "make inferences about a larger population's characteristics/decisions"; and "reconstruct an event or set of events."
6 Bennett and Checkel (2015) differentiate between two forms of process tracing, theory development and theory testing, Beach and Pedersen (2019) between three forms thereof. Due to space constraints, it is not possible to link the expert-interview method to these different process-tracing forms here.
7 Naturally, the longer they are out of the organization in question and the greater the distance to actual decision-making processes, the harder it is to consider these experts "insiders."
8 For a thorough recent discussion of how ethical considerations affect data quality and research results in conflict zones, see Parkinson (2022).
9 More information in Blair and Imai (2012) and Lavrakas (2008); for the application in IR research, see Dietrich, Hardt, and Swedlund (2021).
10 See for instance the controversy about the Data Access and Production Transparency (DA-RT) requirements (e.g., Jacobs et al. 2021; Tripp 2018).

References

Aberbach, Joel D., and Bert A. Rockman. 2002. "Conducting and Coding Elite Interviews." *Political Science & Politics* 35(4): 673–76.

Anderl, Felix. 2015. "Review: Andrew Bennett & Jeffrey T. Checkel (Eds.) (2015). Process Tracing: From Metaphor to Analytic Tool." *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* 16(3). Retrieved September 25, 2019 (http://www.qualitative-research.net/index.php/fqs/article/view/2456).

Angrist, Joshua D., and Jörn-Steffen Pischke. 2010. "The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics." *Journal of Economic Perspectives* 24(2): 3–30.

Archibald, Mandy M., Rachel C. Ambagtsheer, Mavourneen G. Casey, and Michael Lawless. 2019. "Using Zoom Videoconferencing for Qualitative Data Collection: Perceptions and Experiences of Researchers and Participants." *International Journal of Qualitative Methods* 18 1–8.

Autesserre, Séverine. 2014. *Peaceland: Conflict Resolution and the Everyday Politics of International Intervention*. New York: Cambridge University Press.

van Baalen, Sebastian. 2018. "'Google Wants to Know Your Location': The Ethical Challenges of Fieldwork in the Digital Age." *Research Ethics* 14(4): 1–17.

Balcells, Laia, and Patricia Justino. 2014. "Bridging Micro and Macro Approaches on Civil Wars and Political Violence: Issues, Challenges, and the Way Forward." *Journal of Conflict Resolution* 58(8): 1343–59.

Banerjee, Abhijit V., and Esther Duflo. 2009. "The Experimental Approach to Development Economics." *Annual Review of Economics* 1(1): 151–78.

Beach, Derek, and Rasmus Brun Pedersen. 2019. *Process-Tracing Methods. Foundations and Guidelines*. 2d ed. Ann Arbor: University of Michigan Press.

Bennett, Andrew, and Jeffrey T. Checkel. 2015. "Process Tracing: From Philosophical Roots to Best Practices." In *Process Tracing: From Metaphor to Analytic Tool*, eds. Andrew Bennett and Jeffrey T. Checkel, 3–37. Cambridge/New York: Cambridge University Press.

Berry, Jeffrey M. 2002. "Validity and Reliability Issues in Elite Interviewing." *Political Science & Politics* 35(4): 679–82.

Blackwell, Matthew, and Adam N. Glynn. 2018. "How to Make Causal Inferences with Time-Series Cross-Sectional Data under Selection on Observables." *American Political Science Review* 112(4): 1067–82.

Blair, Graeme, and Kosuke Imai. 2012. "Statistical Analysis of List Experiments." *Political Analysis* 20(1): 47–77.

BTI. 2021. "Bertelsmann Transformation Index (BTI)." Gütersloh: Bertelsmann Foundation. Retrieved August 10, 2021 (https://www.bti-project.org/en/home.html).

Clark, Janine A. 2006. "Field Research Methods in the Middle East." *PS: Political Science & Politics* 39(3): 417–24.

Collins, Harry, and Robert Evans. 2007. *Rethinking Expertise*. Chicago: University of Chicago Press.

Cramer, Katherine J. 2016. *The Politics of Resentment: Rural Consciousness in Wisconsin and the Rise of Scott Walker*. Chicago: University of Chicago Press.

Denzin, Norman K. 1978. *Sociological Methods: A Sourcebook*. New York: McGraw-Hill.

Dexter, Lewis Anthony. 2006 [1970]. *Elite and Specialized Interviewing*. Colchester: ECPR Press.

Dietrich, Simone, Heidi Hardt, and Haley J. Swedlund. 2021. "How to Make Elite Experiments Work in International Relations." *European Journal of International Relations* 27(2): 596–621.

Dorussen, Han, Hartmut Lenz, and Spyros Blavoukos. 2005. "Assessing the Reliability and Validity of Expert Interviews." *European Union Politics* 6(3): 315–37.

Driscoll, Jesse. 2021. *Doing Global Fieldwork: A Social Scientist's Guide to Mixed-Methods Research Far from Home*. New York: Columbia University Press.

Dunning, Thad. 2012. *Natural Experiments in the Social Sciences: A Design-Based Approach*. Cambridge: Cambridge University Press.

——. 2015. "Improving Process Tracing." In *Process Tracing: From Metaphor to Analytic Tool*, ed. Andrew Bennett and Jeffrey T. Checkel, 211–236. New York: Cambridge University Press.

Elman, Colin, and Colleen Dougherty Burton. 2016. "Research Cycles: Adding More Substance to the Spin." *Perspectives on Politics* 14(4): 1067–70.

Elman, Colin, John Gerring, and James Mahoney. 2016. "Case Study Research: Putting the Quant into the Qual." *Sociological Methods & Research* 45(3): 375–91.

Elster, Jon. 1998. "A Plea for Mechanisms." In *Social Mechanisms: An Analytical Approach to Social Theory*, ed. Peter Hedström and Richard Swedberg, 45–73. New York: Cambridge University Press.

Fu, Diana, and Erica S. Simmons. 2021. "Ethnographic Approaches to Contentious Politics: The What, How, and Why." *Comparative Political Studies* 54(10): 1695–721.

Fujii, Lee Ann. 2010. "Shades of Truth and Lies: Interpreting Testimonies of War and Violence." *Journal of Peace Research* 47(2): 231–41.

——. 2017. *Interviewing in Social Science Research: A Relational Approach*. Abingdon: Routledge.

George, Alexander L., and Andrew Bennett. 2005. *Case Studies and Theory Development in the Social Sciences*. Cambridge, MA: MIT Press.

Gerring, John. 2017. *Case Study Research: Principles and Practices*. 2d ed. New York: Cambridge University Press.

Glynn, Adam N., and Nahomi Ichino. 2015. "Using Qualitative Information to Improve Causal Inference." *American Journal of Political Science* 59(4): 1055–71.

Goldstein, Kenneth. 2002. "Getting in the Door: Sampling and Completing Elite Interviews." *Political Science & Politics* 35(4): 669–72.

Hafner-Burton, Emilie M., D. Alex Hughes, and David G. Victor. 2013. "The Cognitive Revolution and the Political Psychology of Elite Decision Making." *Perspectives on Politics* 11(2): 368–86.

Hanson, Robin *et al.* 2011. *What's Wrong with Expert Predictions*. Washington, DC: Cato Institute. (https://www.cato-unbound.org/issues/july-2011/whats-wrong-expert-predictions/).

Hardt, Heidi. 2018. *NATO's Lessons in Crisis: Institutional Memory in International Organizations*. New York: Oxford University Press.

Heckathorn, Douglas D., and Christopher J. Cameron. 2017. "Network Sampling: From Snowball and

Multiplicity to Respondent-Driven Sampling." *Annual Review of Sociology* 43(1): 101–19.

Hegre, Håvard *et al.* 2021. "ViEWS2020: Revising and Evaluating the ViEWS Political Violence Early-Warning System." *Journal of Peace Research* 58(3): 599–611.

Irgil, Ezgi, Anne-Kathrin Kreft, Myunghee Lee, Charmaine N. Willis, and Kelebogile Zvobgo. 2021. "Field Research: A Graduate Student's Guide." *International Studies Review* 23(4): 1495–517.

Jacobs, Alan M. *et al.* 2021. "The Qualitative Transparency Deliberations: Insights and Implications." *Perspectives on Politics* 19(1): 171–208.

Kabeer, Naila. 2019. "Randomized Control Trials and Qualitative Evaluations of a Multifaceted Programme for Women in Extreme Poverty: Empirical Findings and Methodological Reflections." *Journal of Human Development and Capabilities* 20(2): 197–217.

Krause, Jana. 2021. "The Ethics of Ethnographic Methods in Conflict Zones." *Journal of Peace Research* 58(3): 329–41.

Lareau, Annette. 2021. *Listening to People: A Practical Guide to Interviewing, Participant Observation, Data Analysis, and Writing It All Up*. Chicago: University of Chicago Press.

Lavrakas, Paul J. 2008. "List-Experiment Technique." In *Encyclopedia of Survey Research Methods*, ed. Paul J. Lavrakas, 433–35. Thousand Oaks: Sage Publications.

Lo Iacono, Valeria, Paul Symonds, and David H.K. Brown. 2016. "Skype as a Tool for Qualitative Research Interviews." *Sociological Research Online* 21(2): 103–17.

Lundgren, Magnus. 2020. "Causal Mechanisms in Civil War Mediation: Evidence from Syria." *European Journal of International Relations* 26(1): 209–35.

Lyall, Jason. 2015. "Process Tracing, Causal Inference and, Civil War." In *Process Tracing: From Metaphor to Analytic Tool*, ed. Andrew Bennett and Jeffrey T. Checkel, 186–208. New York: Cambridge University Press

Maestas, Cherie. 2018. "Expert Surveys as a Measurement Tool: Challenges and New Frontiers." In *The Oxford Handbook of Polling and Survey Methods*, ed. Lonna Rae Atkeson and R. Michael Alvarez, 583–608. Oxford: Oxford University Press.

Meyer, Christoph O., Chiara De Franco, and Florian Otto. 2019. *Warning about War: Conflict, Persuasion and Foreign Policy*. Cambridge: Cambridge University Press.

Mwambari, David. 2019. "Local Positionality in the Production of Knowledge in Northern Uganda." *International Journal of Qualitative Methods* 18:1–12.

Mwambari, David, Andrea Purdeková, and Aymar Nyenyezi Bisoka. 2021. "Covid-19 and Research in

Conflict-Affected Contexts: Distanced Methods and the Digitalisation of Suffering." *Qualitative Research*. OnlineFirst. https://doi.org/10.1177/1468794121999014.

Orne, Martin T. 1962. "On the Social Psychology of the Psychological Experiment: With Particular Reference to Demand Characteristics and Their Implications." *American Psychologist* 17(11): 776–83.

Pakulski, Jan. 2008. "Elite Theory." In *International Encyclopedia of the Social Sciences*, ed. William A. Darity, 562–564. Detroit: Macmillan Reference USA.

Parkinson, Sarah E. 2022. "(Dis)Courtesy Bias: 'Methodological Cognates,' Data Validity, and Ethics in Violence-Adjacent Research." *Comparative Political Studies* 55(3): 420–50.

Peabody, Robert L. *et al.* 1990. "Interviewing Political Elites." *PS: Political Science & Politics* 23(3): 451–55.

Pepinsky, Thomas B. 2019. "The Return of the Single-Country Study." *Annual Review of Political Science* 22(1): 187–203.

Pouliot, Vincent. 2015. "Practice Tracing." In *Process Tracing: From Metaphor to Analytic Tool*, ed. Andrew Bennett and Jeffrey T. Checkel, 237–59. New York: Cambridge University Press.

Rivera, Sharon Werning, Polina M. Kozyreva, and Eduard G. Sarovskii. 2002. "Interviewing Political Elites: Lessons from Russia." *PS: Political Science & Politics* 35(4): 683–88.

Schatz, Edward, ed. 2009. *Political Ethnography: What Immersion Contributes to the Study of Power*. Chicago: University of Chicago Press.

Schedler, Andreas. 2012. "Judgment and Measurement in Political Science." *Perspectives on Politics* 10(1): 21–36.

Schirmer, Swantje. 2021. "'Spotlight On…' Virtual Field Research." Hamburg: German Institute for Global and Area Studies (GIGA). Retrieved August 20, 2021. (https://www.giga-hamburg.de/en/news/news-doctoral-programme-spotlight-on-virtual-field-research/).

Shesterinina, Anastasia. 2016. "Collective Threat Framing and Mobilization in Civil War." *American Political Science Review* 110(3): 411–27.

von Soest, Christian. 2007. "How Does Neopatrimonialism Affect the African State? The Case of Tax Collection in Zambia." *Journal of Modern African Studies* 45(4): 621–45.

——. 2009. *The African State and Its Revenues. How Politics Influences Tax Collection in Zambia and Botswana*. Baden-Baden: Nomos.

Swedlund, Haley J. 2017. "Can Foreign Aid Donors Credibly Threaten to Suspend Aid? Evidence from a

Cross-National Survey of Donor Officials." *Review of International Political Economy* 24(3): 454–96.

Tansey, Oisín. 2007. "Process Tracing and Elite Interviewing: A Case for Non-Probability Sampling." *PS: Political Science & Politics* 40(4): 765–72.

Tetlock, Philip E. 2005. *Expert Political Judgment*. Princeton: Princeton University Press.

Tourangeau, Roger, ed. 2014. *Hard-to-Survey Populations*. Cambridge: Cambridge University Press.

Trachtenberg, Marc. 2006. *The Craft of International History: A Guide to Method*. Princeton: Princeton University Press.

Tripp, Aili Mari. 2018. "Transparency and Integrity in Conducting Field Research on Politics in Challenging Contexts." *Perspectives on Politics* 16(3): 728–38.

de Vries, Lotje, and Tim Glawion. 2021. "Studying Insecurity from Relative Safety — Dealing with Methodological Blind Spots." *Qualitative Research*. OnlineFirst. https://doi.org/10.1177/14687941211061061.

Wackenhut, Arne F. 2018. "Ethical Considerations and Dilemmas Before, During and After Fieldwork in Less-Democratic Contexts: Some Reflections from Post-Uprising Egypt." *American Sociologist* 49(2): 242–57.

Wedeen, Lisa. 2010. "Reflections on Ethnographic Work in Political Science." *Annual Review of Political Science* 13(1): 255–72.

Wood, Elisabeth Jean. 2006. "The Ethical Challenges of Field Research in Conflict Zones." *Qualitative Sociology* 29(3): 373–86.