

# Advances in Survey Methods for the Developing World

Noam Lupu and Kristin Michelitch

Department of Political Science, Vanderbilt University, Nashville, Tennessee 37203, USA;  
email: noam.lupu@vanderbilt.edu, kristin.michelitch@vanderbilt.edu

Annu. Rev. Political Sci. 2018. 21:195–214

First published as a Review in Advance on  
January 31, 2018

The *Annual Review of Political Science* is online at  
polisci.annualreviews.org

<https://doi.org/10.1146/annurev-polisci-052115-021432>

Copyright © 2018 by Annual Reviews.  
All rights reserved

## Keywords

survey research, developing countries, survey sampling, data collection, research ethics, research transparency

## Abstract

Political scientists are fielding more and more surveys in the developing world. Yet, most survey research methodology derives from experiences in developed countries. Researchers working in the developing world often confront very different challenges to collecting high-quality data. Census data may be unreliable or outdated, enumerators may shirk, political topics may be sensitive, and respondents may be unaccustomed to and uncomfortable with an interview format. In this article, we review both published methodological research and the best practices of scholars based on an original expert survey of survey researchers. We characterize the state of the field and provide insights about the range of available options for implementing surveys in the developing world. We examine and assess innovations across many aspects of survey implementation, including sampling, enumeration, data collection, ethical considerations, and reporting. We also offer suggestions for future methodological inquiry and for greater research transparency.



### ANNUAL REVIEWS Further

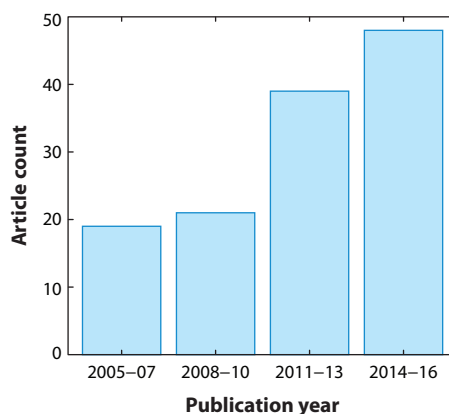
Click here to view this article's  
online features:

- Download figures as PPT slides
- Navigate linked references
- Download citations
- Explore related articles
- Search keywords

## INTRODUCTION

Political scientists are increasingly conducting survey research in the developing world. Cross-national barometer surveys have sprung up in every developing region (Heath et al. 2005, Smith 2015), and developing countries are increasingly represented on global projects such as the Comparative Study of Electoral Systems and the World Values Survey. Individual researchers are also increasingly fielding original surveys across the developing world,<sup>1</sup> no doubt buoyed by decreasing costs and improvements in local capacity (see Zechmeister & Seligson 2012). Indeed, the number of articles published in the top political science journals in which authors field original surveys from the developing world has ballooned in the past decade (see **Figure 1**).<sup>2</sup>

Yet, most survey research methodology derives from experiences in developed countries, particularly the United States. Researchers working in the developing world often confront very different challenges to collecting high-quality data.<sup>3</sup> For example, how do researchers sample in a country without reliable and up-to-date census data? How do we elicit opinions on national politics in closed societies or dangerous security environments? How do we deal with cultural



**Figure 1**

Political science publications using original surveys from the developing world. Figures are the absolute number of articles using an original survey from a developing country published between 2005 and 2016 in *American Journal of Political Science*, *American Political Science Review*, *British Journal of Political Science*, *Comparative Political Studies*, *Electoral Studies*, *Journal of Politics*, *Political Behavior*, *Public Opinion Quarterly*, or *World Politics*. We classify as “developing” any country other than Andorra, Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Israel, Italy, Japan, Liechtenstein, Luxemburg, Malta, Monaco, the Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, the United Kingdom, and the United States.

<sup>1</sup>We classify as “developing” any country other than Andorra, Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Israel, Italy, Japan, Liechtenstein, Luxemburg, Malta, Monaco, the Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, the United Kingdom, and the United States. We have reservations about the “developing/developed” terminology, but settled on it because it is the most widely used in political science and, in our estimation, the least offensive way to refer to the set of countries we have in mind.

<sup>2</sup>In this classification, and throughout this article, we include both mass and elite surveys. Following Groves et al. (2009), we define a survey as a large-N, systematic method of collecting information about a population from a subset.

<sup>3</sup>A recent exception is the compilation of guidelines in SRC (2016), as are the related edited volumes on multinational, multiregional, and multicultural (3MC) surveys (Harkness et al. 2010, Johnson et al. 2018). Although pitched at cross-national survey projects, SRC (2016) is also a useful reference for some of the general challenges of fielding surveys in developing countries.

norms regarding gender roles or hospitality in face-to-face surveys? How should we monitor enumerators? What ethical challenges do we need to consider?<sup>4</sup>

In this review article, we rely on three sources of information about these issues. As with most review articles, a first source is the published research on the topic—in our case, research on survey methodology that deals with questions of particular relevance in developing contexts. Existing methodological research has produced many insights. However, published work has yet to catch up with the new challenges arising as scholars increasingly conduct surveys in the developing world. Many pressing challenges have simply not been taken up by methods scholarship.


Where there is little or no guidance in published studies of survey methodology, we turn instead to best practices. Although scholars may have yet to publish systematic studies on some of the methodological issues that arise when conducting surveys in the developing world, researchers fielding surveys regularly find ways to resolve them. To collect this information about survey research in practice, our research assistants combed through the methodological information available about the major barometers (the Afrobarometer, AmericasBarometer, Arab Barometer, Asian Barometer, Central and Eastern Eurobarometer, and Latinobarometer) as well as every journal article containing an original survey conducted in a developing country that was published between 2005 and 2016 in *American Journal of Political Science*, *American Political Science Review*, *British Journal of Political Science*, *Comparative Political Studies*, *Electoral Studies*, *Journal of Politics*, *Political Behavior*, *Public Opinion Quarterly*, or *World Politics*. They culled relevant information about the method of survey implementation from the published articles themselves, from methodological or online appendices, and from any other publicly available resource. As it turned out, these sources reported precious little detailed information about how each survey was implemented; indeed, very few provided any information beyond a brief description of sample design and, occasionally, the response rate. We return in a later section to this issue of reporting and transparency in survey research.

Given the thinness of methodological literature and paucity of information regarding how published survey data from the developing world are collected, we decided to field our own survey of experts as a third source. We asked the top researchers conducting surveys in the developing world how they overcame key challenges.

This review article draws from all three of these sources to characterize the state of the field and to provide insights about the range of options available for implementing surveys in the developing world. We examine innovations across many aspects of survey implementation, including sampling, enumerator effects, eliciting high-quality responses, ethical considerations, and reporting. We close by offering suggestions for future methodological inquiry.

## THE EXPERT SURVEY

We fielded an online survey, via Qualtrics, of experts who had conducted an original survey in a developing country. Our complete questionnaire and data set are available as **Supplemental Materials**. We invited scholars who either were principal investigators on one of the regional barometers or had published an article in *American Journal of Political Science*, *American Political Science Review*, *Comparative Political Studies*, *Journal of Politics*, *Public Opinion Quarterly*, or *World Politics* between January 2010 and October 2015 that analyzed an original survey in a developing country. We excluded authors of published articles that analyzed barometer data who were not principal investigators on those barometers. We also excluded ourselves.

 **Supplemental Material**

<sup>4</sup>Some of these challenges also arise when conducting surveys of hard-to-reach populations in more-developed countries (see, e.g., Tourangeau et al. 2014).

Experts were first invited to participate in December 2015, reminders were sent after several weeks, and the survey was closed in March 2016. Of the 175 experts we invited to participate, 80 responded to the survey, a response rate of 46%. For respondents who completed the survey in one sitting, it took, on average, 24 minutes. As is typical of online surveys (Peytchev 2009), 30 of our 80 respondents (38%) terminated the survey early. All of the averages and proportions we report for a particular question in the expert survey are based on the total number of responses to that question. Our average respondent was 40 years old, and of those who reported their gender, 67% were men.

We asked each respondent a series of questions about the most recent original survey with some political content that they fielded in a developing country. Among the most recent surveys they had conducted, 41% were fielded in Latin America, 26% in sub-Saharan Africa, 17% in Asia, 9% in the post-Soviet bloc, and 7% in the Middle East and North Africa. Unsurprisingly, the vast majority of these surveys were face-to-face (85%); most of the rest were over the phone, with a handful of online surveys. Among the surveys conducted face-to-face, 51% recorded responses on an electronic device. Most of the surveys used national samples (58%), and only a small subset were panels (13%). The average sample size was 1,800 observations (excluding two unusually large samples of more than 10,000). On average, the length of the survey was 35 minutes (excluding two outlier experts who reported their surveys taking 175 minutes and 300 minutes). Surveys were most commonly funded through external government grants (52%), but some were also supported by university grants (30%) or scholars' research accounts (20%).

Most of our respondents (68%) had hired a local survey firm or academic institute to field their survey. Among the rest, one-third reported that there were no survey firms or academic institutes in their target country capable of fielding their survey. Many of the surveys employed visual or audio aids (43%), and a fairly large number used behavioral tasks (22%). Impressively, the vast majority of our expert respondents reported including some kind of experiment in their survey (68%), a testament to the growth of survey experiments (Druckman et al. 2006, Mutz 2011).

## SAMPLING

Most academic surveys, especially mass surveys, in the developing world continue to be conducted face-to-face, as we found in the methodological literature, substantive articles, and the expert survey. In most cases, internet and phone penetration are still not high enough to achieve representative samples. Even where there is relatively high usage of phones, telephone sampling frames are often systematically incomplete. And in much of the developing world, cell phones—which may not be listed in telephone directories or officially registered—are far more prevalent than land lines. Address-based sampling (Iannacchione 2011), which requires reliable address directories, tends also to be infeasible in developing contexts. Instead, most scholars conduct face-to-face surveys and use some form of multistage area sampling that proceeds from geographic units to households to individuals (see, e.g., Kalton 1983, Lohr 2010). For convenience and cost savings, scholars typically use cluster sampling at some stage. In our expert survey, the vast majority of our respondents (77%) said their most recent survey involved some type of area sampling. The next-largest category was convenience sampling at 11%.

Multistage cluster sampling requires some information about how to divide the target population, both geographically and numerically. The respondents to our expert survey who used this sampling method typically used census information or maps to identify geographic units and census data or other government agency data to determine the population size in each unit. Then, once enumerators arrived at the enumeration area, they typically selected a block and approached every

*n*th household, often excluding potential subjects residing in institutions like hospitals or military barracks. This procedure is difficult to employ in rural settings where dwellings are not organized into blocks, so some of our respondents—and many published surveys—instruct enumerators to select households along a random walk. Of course, human beings are notoriously bad at randomizing, and recent research shows that enumerators' instructions on conducting a random walk systematically affect both selection probabilities and survey results (Bauer 2014, 2016). Although less common in more-developed contexts, quotas (typically gender and age groups) are regularly employed in mass surveys in the developing world to ensure demographic balance. In our expert survey, 59% of the face-to-face surveys employed them. In the absence of quotas, researchers typically construct post-stratification weights, with raking being the most widely accepted method (Kalton & Flores-Cervantes 2003).

Across surveys, there is a great deal of variation in the number of times a selected household is visited before being replaced. Among the surveys conducted by our expert respondents, just over two-thirds revisited a selected household at least once. One expert reported doing up to five revisits. In principle, revisits—or callbacks in phone surveys—should increase response rates and reduce the potential for nonresponse bias. But revisits can also be costly: They require that enumerators, who may not reside in or near the enumeration area, remain in the area or return to it multiple times. We know of no studies focused on the effects of revisits in developing contexts, but recent research suggests there is little gain from repeated costly attempts to contact potential respondents (Legleye et al. 2013). Another approach is to make contact with potential participants in advance and set up appointments for later interviews with the enumeration team. Roughly one in five of our expert respondents said they did this in their most recent survey. Yet, there is still little empirical research demonstrating the benefits of these approaches to nonresponse and representativeness.

A common challenge in applying area sampling techniques in the developing world is that census data may be unreliable, outdated, or unavailable. This challenge may be especially large among populations that are informally settled, where rates of violence and migration are high, or in nondemocratic contexts where government information is not publicly available. A small number of our experts and a small number of published articles in political science use alternatives involving satellite and global positioning system (GPS) technology. In the absence of reliable census data on urban migrant settlements in China, Landry & Shen (2005) developed a method for using GPS coordinates to select enumeration areas (see also Landry 2010). Driscoll & Lidow (2014) used high-resolution satellite images and remote sensing technology to gather population information in Somalia in the absence of up-to-date government data (see also Montalvo et al. 2018). These techniques, which are being developed especially rapidly in research on public health and development, will no doubt become more common in political science as these technologies continue to improve.

In face-to-face surveys, enumerators are typically instructed to keep track of every household they visit so that the researcher can calculate the survey response rate. As defined by most scholars, the response rate is the number of completed interviews divided by the number of eligible individuals selected into the sample (Groves et al. 2009). The gold standard is to employ the final disposition codes set out by the American Association for Public Opinion Research (AAPOR 2016) and use those frequencies to calculate a response rate—typically the minimum response rate, RR1. One issue in calculating response rates is how to deal with cases of unknown eligibility. For instance, if an individual answers the door and immediately turns the enumerator away, it may be impossible to know whether that individual met the eligibility criteria for the survey. Again, AAPOR offers researchers guidelines and several options—each making different assumptions—for calculating the likely proportion of eligible subjects (Smith 2009b).

The methodological literature identifies various factors that contribute to an individual's decision to participate in a survey: individual predispositions, contextual factors, the topic of the survey, and the disposition of the enumerator, among others (Banducci & Stevens 2015, Durrant et al. 2010, Groves & Couper 1998). Surveys conducted in the developing world tend to report higher response rates than those conducted in more-developed countries. One reason may be that face-to-face surveys, which are more common in developing contexts, tend to yield higher rates of response than phone or online surveys.<sup>5</sup> A second reason may be that public opinion surveys are simply less common in these contexts and therefore more interesting for respondents. The respondents in our expert survey reported an average response rate in their surveys of 74%. Among the published articles with mass surveys in developing countries that reported a response rate, the average was 69%. For comparison, the 2016 American National Election Study reports a 50% response rate for its face-to-face component and a 44% response rate for the online component.

Moreover, there are broader debates—and quite a lot of hand-wringing—about declining response rates in more-developed contexts and their impact on scholarship (Groves et al. 2002, 2006; Massey & Tourangeau 2013). If nonrespondents were similar to survey respondents, then response rates would be mere curiosities. But most scholars suspect that response rates are associated with some nonresponse bias. The empirical evidence on this, however, is mixed (Groves & Peytcheva 2008), and mostly focuses on more-developed countries. We still know little about how response rates vary in the developing world across space and time, what techniques help increase participation, and the extent to which nonresponse induces bias.

## ENUMERATOR EFFECTS

Given the prevalence of face-to-face surveys in developing contexts, enumerators often play a crucial role in data collection. In countries with well-developed polling operations, enumerators are often professionals who make their living enumerating surveys. Where surveys are less frequent, firms and nongovernmental organizations (NGOs) regularly recruit university students willing to do part-time work. For the surveys conducted by our expert respondents, most of the enumerators had been recruited by some member of the research team, either one of the principal investigators or a research assistant. In one-third of the cases, the contracted survey organization recruited the enumerators.

The role of enumerators in shaping survey outcomes and introducing error in face-to-face surveys has been receiving increasing scholarly attention. Bischooping & Schuman's (1992) classic pen experiment in Nicaragua showed that even the seemingly minor choices of enumerators—in their case, pen color—can have significant effects on responses. More recently, scholars have documented a range of interviewer effects (see Schaeffer et al. 2010, West & Blom 2017), stemming primarily from the effects of observable interviewer characteristics. The gender (Benstead 2014b, Flores-Macías & Lawson 2008, Liu & Stainback 2013), ethnicity (Adida et al. 2016), and apparent religiosity (Benstead 2014a, Blaydes & Gillum 2013) of the enumerator all seem to especially influence survey responses. For the most part, the particular interviewer attribute seems to affect the response in intuitive ways: religiosity affecting questions about religiosity among coreligionists, gender affecting questions about gender issues, and so on. Survey researchers should consider these effects when recruiting enumerators; for certain studies and in certain contexts, it may be important to ensure that enumerators and respondents are matched in gender, ethnicity, or religion.

<sup>5</sup>However, this may be at least partly an artifact of the reliance on enumerator reports. Enumerators, often paid on a per-interview basis, have little incentive to waste time tracking every attempted contact. As a result, response rates based on enumerator reports may well underestimate the number of contacts, inflating the response rate.



In general, the literature underscores that interviewer characteristics seem to especially affect responses to attitudinal, sensitive, complex, and open-ended questions—although they sometimes also affect responses to factual questions (see Davis & Silver 2003). There is also some evidence that interviewers' own opinions affect survey responses (Himelein 2016). An enumerator's gender, level of experience, and personality also appear to affect both responses and the willingness of potential respondents to participate (Blom et al. 2007, Durrant et al. 2010, Olson & Peytchev 2007, West & Blom 2017, West & Olson 2010). And interviewer characteristics—particularly experience—seem also to affect paradata such as interviewer observations of the respondent, household, or neighborhood (Lupia 2015, Sinibaldi et al. 2013, West & Kreuter 2013).<sup>6</sup> Interviewers probably affect survey error in every study, but how they affect it varies by context and survey design. As more and more surveys are conducted in the developing world, we ought to continue to expand our understanding of these effects and the circumstances that exacerbate or attenuate them.

Interviewer effects are difficult to identify because interviewers are rarely randomly paired with respondents. Instead, most studies use hierarchical models that nest respondents within interviewers and control for interviewer characteristics. Still, we know that certain types of interviewers are more likely to be assigned to areas with certain types of respondents; for instance, male interviewers and those with more experience are often sent to neighborhoods with higher crime rates.

Another major challenge in enumerator-administered surveys is monitoring. Often paid per completed interview, enumerators have strong financial incentives to fabricate interviews. If they are being sent into high-crime or remote enumeration areas, they may also have personal incentives to shirk. They may shorten an interview by skipping some questions and making up the responses, or they may fabricate an entire interview that never took place. In quota samples, enumerators may lie about a respondent's age or gender to fill their daily quota. Middle-aged women tend to be easier to find and more willing to participate, so enumerators may interview them and claim they had interviewed harder-to-reach demographic groups like young men. Nearly one-third of our expert respondents reported having thrown out some interviews in their data because of enumerator dishonesty.

A recent method for detecting fabricated data caused an uproar with the claim that 26% of publicly available surveys in the developing world contain significant amounts of fabricated data (Kuriakose & Robbins 2016; see also Gelman 2016 and the response by Robbins 2016, as well as Koczela et al. 2015, Robbins 2018, Slomczynski et al. 2017). The method works by identifying near-matching observations in survey data sets, under the assumption that enumerators fabricating interviews are likely to generate responses that are unusually similar. Although the specific claim about the extent of the problem is under debate (see, e.g., Simmons et al. 2016), researchers have also improved their monitoring efforts.

Survey researchers regularly do some kind of respondent follow-up—either in person or over the phone—to validate some subset of interviews. Roughly half of our expert respondents reported having done this as part of their most recent survey. Recent advances in enumerator monitoring have been made possible by the use of electronic devices such as smartphones or tablets to conduct interviews; just over half of our expert respondents conducted their most recent survey on electronic devices. Unlike the traditional pencil and paper, electronic devices allow audio recording of enumerators reading questions aloud, identify the location of interviews using GPS, and track the duration of interviews and individual items (e.g., Gomila et al. 2017). Cross-national barometer projects such as the AmericasBarometer have already developed extensive quality-control

<sup>6</sup>Among the surveys conducted by our expert respondents, nearly half solicited this type of interviewer-observed paradata.

protocols that include audio recording, enumerator photo captures, GPS tracking, and other electronic tools (Montalvo et al. 2018).

## ELICITING HIGH-QUALITY RESPONSES

Once enumerators select respondents and follow protocol for interviewing them, it can also be challenging to elicit high-quality responses. Enumerators are typically trained to find a quiet and private place in which to conduct an interview, but this can be difficult or impossible in some developing contexts. Without privacy from bystanders, respondents may self-censor their responses, although the limited empirical evidence on this score is mixed (Diop et al. 2015, Smith 1997). Where communal living is common, private spaces may not exist. Even if they do, communities may have norms against sitting with a stranger in private. Taboos against taking an enumerator of the opposite sex into a private setting are especially common, but sometimes also apply to noncoethnics. In some contexts, it is impossible to conduct an interview without the (often male) head of the household being present. Moreover, where crime rates are high, respondents may be reluctant to allow an enumerator into their home, preferring to conduct the interview outside, in the doorway, or through a window. According to the most recent round of the Afrobarometer, for example, bystanders were present during 44% of interviews. In our expert survey, the most common bystanders named were husbands.

One way researchers mitigate some of these situations is by ensuring that interviews are conducted by cogender or coethnic enumerators (e.g., Corstange 2009). Bleck & Michelitch (2017) had mixed-gender enumerator teams simultaneously interview one man and one woman in each household in rural Mali to facilitate privacy for female respondents. Another option is to make some portion of the interview self-administered and ensure the respondent's privacy during that section (see, e.g., Chauchard 2013). Where bystanders are unavoidable, researchers can also ask enumerators to record that a bystander was present—and who it was—and account for potential bystander influence in the analysis.

Respondents themselves may have difficulty responding to enumerator questions. In the developing world, lower levels of literacy and numeracy, multiple languages, and the existence of local dialects and oral languages can complicate enumeration. Low literacy and numeracy can restrict the types of questions and visual aids researchers can utilize. Respondents with little formal education may find it difficult, for instance, to think in terms of proportions, and may be unaccustomed to the exam-like style of closed-ended survey questions; their discomfort may increase systematic item nonresponse or satisficing (see Krosnick 1991). They may also find it difficult to self-administer a portion of the survey.

Low literacy may also make it ineffective for researchers to embed experiments with lengthy texts into their survey, and researchers may instead need to consider visual aids, cartoons, or videos. Even when only some portion of the target population is less literate, its likely correlation with other characteristics and political attitudes means that researchers ought to consider designing their questionnaire to accommodate these respondents.

Where formal schooling and surveys are less widespread, respondents may be unaccustomed to answering questions, especially on political topics. They may find the interview format very awkward and the notion of closed-ended responses foreign. Transitions between different question formats may be challenging. In our expert survey, 35% of researchers reported that some respondents had trouble discussing politics, naming women as the group with most difficulty. It can be useful to have enumerators identify whether individual respondents seemed to have more or less difficulty with the survey. In the most recent Afrobarometer round, enumerators reported that 52% of respondents had difficulty answering questions. In some of the poorest countries,



however, the vast majority of respondents did: 95% in Guinea, 93% in Mali, 91% in Niger, 86% in Senegal, and 81% in Benin and Gabon. Simplifying questionnaires can surely help with these issues, but researchers should also recruit enumerators who can patiently assist respondents. As Paluck (2009, p. 45) relates, “it is an important methodological and even ethical consideration to compose a research team that will invite the confidence and frankness of the whole range of participants in the sample.”

More generally, survey researchers in developing contexts often have to field their studies in multiple languages, even within the same country. Among multilingual populations, there may be benefits to offering respondents multiple interview-language options (see Lee & Perez 2014). Among our expert respondents, 67% reported translating their questionnaire, half to only one language and the rest ranging from 2 to 10 languages. Most experts translated their instruments themselves, with feedback from local experts. High-quality translations are especially important in instances when a single study is conducted in multiple languages, as in the case of most cross-national projects. Even when rigorous translation procedures are implemented, it can be very difficult to generate equivalent items across languages (e.g., Davidov & De Beuckelaer 2010, Heath et al. 2009; for an example focused on party identification, see Johnston 2006). Since language is often correlated with characteristics such as ethnicity, religion, or class, nonequivalence can bias inferences: An apparent ethnic or class difference in responses on an item may instead be an artifact of nonequivalence across translations.

Harkness (2007) and Smith (2010) provide overviews of issues in survey translation. Scholars have proposed a number of alternative translation techniques, including back-translation (Brislin 1970), team translation (Harkness 1999), or some combination (McGorry 2000). However, we are not aware of any studies that empirically assess the relative success of these different methods in generating more equivalent instruments or responses.

In some instances, researchers may allow enumerators to translate a questionnaire on the spot. Only about 10% of our expert respondents had done this. Advance translations probably reduce measurement error, but they may not always be feasible. In some developing contexts, local languages and dialects may be oral. In some parts of sub-Saharan Africa, mother tongues are rarely used in writing, for instance, and some local dialects of indigenous or creole languages in Latin America are also oral. It may be impossible to translate a survey instrument into these local languages, or to expect even local enumerators to be able to read it. In these cases, researchers may have no choice but to rely on local enumerators to translate the survey on the spot, which entails extra time training enumerators in the intended meaning of each question.

The kinds of questions researchers can ask vary dramatically across contexts. More than half of our expert respondents reported asking sensitive questions on their surveys. Especially in non-democratic contexts, political questions can be very sensitive (Tsai 2010) and can yield unreliable responses. As Chia (2014, p. 384) sums up, “people’s willingness to voice opinions may be a function of their perceptions of the government’s stance, and individuals are likely to refrain from expressing opinions that appear to contradict the authority.” Even respondents in democratic countries may be reluctant to voice their support for certain groups or individuals. Greater political uncertainty and higher rates of crime and violence can make respondents in developing countries more reluctant to be truthful. For instance, asking respondents how they voted in an election seems trivial in more-developed democracies, but the question may be far more sensitive where elections are still new and accompanied by partisan violence. These settings make survey participation itself riskier, forcing researchers to think seriously about ensuring respondent confidentiality. In a face-to-face survey using an electronic device, it is possible for the enumerator to encrypt and lock responses upon completing an interview, ensuring their immediate inaccessibility.

Question sensitivity becomes especially problematic if respondents think that the survey is being sponsored by a political actor. Researchers typically provide enumerators with study introduction scripts that emphasize the academic purpose of their studies. Even so, respondents may falsely believe that surveys asking numerous questions about politics are sponsored by political actors. In the most recent round of the Afrobarometer, which asks respondents at the end of the survey who they think sponsored it, 38% of respondents thought it was their government, a political party, or a politician. The 2016 AmericasBarometer in Ecuador asked a similar question at the end of its survey and found that nearly half of respondents thought the survey was sponsored by the government or another political actor.

Scholars have developed a range of tools to elicit truthful responses on sensitive issues (e.g., Tourangeau & Yan 2007). One common finding is that responses to sensitive questions tend to be more reliable when the survey is self-administered rather than enumerator-administered (e.g., Tourangeau & Smith 1996). Since many surveys in developing countries are administered by enumerators, researchers can make the sensitive portion of the survey self-administered (e.g., Beber et al. 2014), assuming respondents are willing and able to self-administer it. With convincing assurances of the anonymity of their responses, respondents may be more truthful. It also seems useful to ask enumerators to report their impressions of the respondent's truthfulness, although these impressions may be subject to bias.

Among political scientists, experimental approaches to sensitive questions have become more common. These include list experiments (e.g., Corstange 2009, Glynn 2013, Gonzalez-Ocantos et al. 2012, Oliveros 2016), endorsement experiments (e.g., Aronow et al. 2015, Blair & Imai 2012, Lyall et al. 2015), and randomized-response techniques (e.g., Gingerich 2010, Gingerich et al. 2016, Jann et al. 2012). All of these approaches attempt to assure respondents that they can be truthful because neither the enumerator nor the researcher will know how they responded,<sup>7</sup> but they require some sophistication on the part of the respondent and usually also necessitate literacy. Other scholars have used implicit association tests to reveal biases respondents may wish to occlude or may not even recognize (e.g., Clayton 2015, Lowes et al. 2015). But these tests must be done on a computer, which can be distracting and unfamiliar to respondents in poorer contexts.

It can be challenging to generate high-quality survey response data in developing contexts. Researchers have proposed and implemented an impressive variety of resourceful solutions to these challenges. How effective are they, which solutions are more effective than others, and under what circumstances do they work best? Future methodological research ought to address these questions.

## ETHICAL CONSIDERATIONS

Survey research involves human subjects, and most universities—particularly in the United States and Europe—require scholars conducting survey research to secure advance approval from an Institutional Review Board (IRB). Among our expert survey respondents, 78% reported receiving IRB approval, exemption, or other authorization at their home institution. Of those working with collaborators, 70% reported knowing that their collaborators had also received IRB approval. They reported that the approval process took between 1 and 20 weeks, with the mode being one month. Many reported difficulties with the process; one-third of the experts rated their IRB process as “difficult” or “very difficult.”

<sup>7</sup>However, the researcher can use new methods to assign probabilities to the response options for each individual (Imai et al. 2015).

A great deal has been written about IRB standards and the ethics of research with human subjects (e.g., Seligson 2008; for an overview of ethical codes in survey research, see Smith 2009a). Researchers face particular challenges when seeking to conduct survey research in developing countries. These obstacles often result from IRB reviewers being unfamiliar with the research site, requirements for in-country research permits, and the pace of change in some developing contexts.

One key challenge is that many university IRB standards are written for clinical research, and IRB personnel are typically unfamiliar with most international contexts, as reported in our expert survey. As a result, IRB reviews often ask researchers planning surveys in developing countries for additional information, slowing the review process (see Zechmeister 2016). One of our respondents, for instance, noted that their review process had been stalled on the issue of consent. The IRB reviewers preferred that survey respondents give written consent, but the researcher felt that written consent forms would put subjects at greater risk, since they would reveal the identity of respondents and could fall into the hands of political actors. In some contexts, moreover, low levels of literacy may make respondents uncomfortable about signing consent forms. Inflexible IRB reviewers with little knowledge of the context in which proposed survey research is sited can slow down the approval process. Universities should consider incorporating in-person or phone interviews with researchers applying for IRB approval of international research.

A common stumbling block with review boards, also emphasized by respondents in our expert survey, centers on incentives. In some cultures, it is simply unacceptable to arrive at a stranger's door without offering a gift. Researchers may also feel obliged to compensate subjects for their time, or use incentives to increase participation rates. On the other hand, there are valid ethical concerns about incentives that are high enough to be coercive. In some contexts, compensation may be interpreted as a bribe for certain responses. Further, offering, for example, a poor laborer a day's wage for a half-hour interview may induce them to respond to sensitive questions (i.e., accept more risk) they might otherwise have preferred to avoid. One recent study in the United States found no effect of higher monetary payments on risk taking (Singer & Couper 2008), but we still know little about how incentives operate in the developing world.<sup>8</sup> In the meantime, researchers should use their contextual knowledge—and the advice of local experts—to evaluate the coercive versus beneficial potential of monetary or in-kind incentives, especially when researching low-income populations or contexts where resources are scarce.

Increasingly, university IRBs are requiring researchers to obtain in-country research permits or seek approval for their study from in-country IRBs. Research review boards are rare in developing countries, and those that do exist tend to focus on medical research exclusively. But many developing countries do require foreign researchers to obtain a government-issued permit to conduct research. Among the experts in our survey, only a handful said they obtained an in-country research permit. Most said such a permit was not required in the country they studied, but several said they did not obtain one even when it was. The process can become complicated when permit-granting agencies ask the researcher to obtain university IRB approval first and the university IRB requires the in-country permit before granting approval.

Securing an in-country research permit can prove challenging. The process can be lengthy, since government agencies issue these types of permits infrequently. Especially in less democratic

<sup>8</sup>There are also other important considerations in evaluating the use of incentives (e.g., Singer et al. 1998). Moreover, researchers ought to consider local practices among other survey researchers. If relatively well-funded international researchers set expectations of remuneration, surveys may become more costly for local researchers.

contexts, governments may seek to block or delay research, particularly on political topics. Non-democratic regimes often impose particularly stringent regulations on survey research (see, e.g., Manion 2010). Bureaucrats in permitting agencies may also expect or demand bribes in order to issue a research permit or speed up the process, putting a foreign researcher in a difficult ethical and legal position. Thus, while there are good reasons to ask researchers to provide some kind of local assessment of their proposed studies with human subjects (see Zechmeister 2016), rigid requirements for official in-country permits or local IRB approval may impede research and raise ethical issues in some contexts. An alternative required by some boards is a written assessment from an in-country researcher.

Conducting human-subjects research in developing countries is often costly and time-consuming and can require adaptation. Repeated pilot studies may be necessary to validate translations, particularly when working in languages that are predominantly oral. Researchers may also need to pretest their questionnaires repeatedly when working with populations less accustomed to responding to surveys or taking tests. The researcher must be able to make changes to questionnaires and other protocols while in the field. Moreover, the political climate on the ground can change quickly, requiring researchers to adjust their research plans.<sup>9</sup> All of these factors make it difficult for researchers to interact with review boards that rightly prefer to use extra caution when evaluating international research. One way to mitigate bureaucratic delay is for researchers to seek IRB approval for the broad contours of a study, before, say, the instrument is finalized, and to submit revisions to an approved protocol from the field. If relatively minor revisions can be reviewed quickly, this approach can both allow reviewers time to thoughtfully evaluate the proposed study and provide room for the researcher to make last-minute changes that are approved quickly.

Although university IRBs typically focus only on protecting respondents, survey researchers working in the developing world also have broader ethical responsibilities to consider (see Fujii 2012). One is protecting enumerators and local collaborators. A research study must consider not only the risk to which respondents are exposed—which is typically minimal in survey research—but the far greater risk to which some enumerators are exposed. Those enumerators being sent to high-crime areas, neighborhoods controlled by gangs, or conflict zones put themselves at risk for the research study. Enumerators working in very poor environments, with little access to food during the workday, or on difficult subjects such as memories of violence, also take on physical and psychological risks (Paluck 2009). It is the researcher's ethical obligation to work to minimize these risks by replacing high-risk enumeration areas, sending enumerators to high-risk areas in teams rather than individually, equipping enumerators with cell phones and battery packs, setting aside time and resources for physical and psychological rest, and working with local communities and organizations to ensure safe passage for enumerators coming into high-risk enumeration areas as outsiders.

## REPORTING SURVEY METHODS

When conducting a survey in the developing world, researchers make many far-from-mundane decisions—about what sampling frame to use, how to select and train enumerators, whether to use paper-and-pencil questionnaires or electronic devices, etc.—that affect the quality of the data. Moreover, many scholars innovate techniques to address the challenges of their particular study

<sup>9</sup>Our own research projects have been disrupted by political events: in Michelitch's case by a coup in Mali (Bleck & Michelitch 2017) and in Lupu's case by the Russian annexation of Crimea (Lupu & Peisakhin 2017).

that could be replicated by future scholars. But we lack a common standard for which aspects of survey research design should be reported.<sup>10</sup> The published literature fails to highlight the impressive range of techniques being employed by researchers to collect high-quality survey data. This omission limits our ability to replicate and to build upon the methods used in previous research.

Among the substantive articles we identified that relied on some survey conducted in the developing world, most described—either in the main text or in an auxiliary appendix—the sampling strategy. Nevertheless, we were unable to find information about the sample design for 4% of the surveys analyzed in these articles. Where we did find information, it was typically vague. Authors often noted that their survey consisted of a “national probability sample of country X” and left it at that. Others characterized their survey as using a “multistage stratified sample” with little or no detail about the strata used and the selection criteria within strata. Particularly in the case of national samples, researchers often have to substitute (usually a small number of) selected sampling units that are too remote, too dangerous, or inaccessible. But only 16% of these studies reported substituting enumeration areas, and of those, only one in five explained how the substitutions were selected. Mass surveys also typically exclude institutionalized individuals—residents of student dormitories, hospitals, or military barracks, for instance—but almost none of the articles mentioned these exclusions. Authors often reported the response rate, although even that was only available for 53% of cases, and typically no information was offered about how nonresponse information was collected or how response rates were calculated.

The regional barometers and cross-national survey projects typically report more methodological information. Still, some of the barometers provide no information about their sample designs. Only the Afrobarometer currently provides information about enumeration area substitutions, and only for some rounds. None of the barometers currently provides information about how nonresponse information was collected.

For nearly one in five of the surveys we reviewed, we were unable to ascertain who conducted the fieldwork: a survey firm, an NGO, or the researchers themselves. Only one in four surveys reported whether a pretest or pilot had been conducted. Nearly half did not report whether the survey was translated. Of those that did report that the survey was translated, only half reported how the survey was translated. We were able to find the full questionnaire for only 20% of the surveys used in these articles. When analyzing the survey data, only one in five articles employed post-stratification weights, and among those, only a handful reported the source of the population data, the specific variables employed, and the weighting method.

Our expert survey revealed a higher level of transparency, although most of it was aspirational. One-third of our respondents reported that the sampling protocol for their most recent survey is available online, while another 65% said they intended to post it. The figures for response rates were nearly identical. Only 22% reported that their most recent survey’s complete data set is available online, but another 71% said they planned to make it available. Among our respondents, 21% reported that the English-language version of their questionnaire is available online, and another 57% said they intended to make it available online. Remarkably, though, 16% of our respondents reported no intention of making their questionnaire available online in the language in which it was administered.

<sup>10</sup> We describe here the information that is currently available online. Although we are certain that researchers would be willing to share additional information upon request, we think the field would benefit from making a set of standard information publicly available online.

In short, survey researchers conducting original surveys in the developing world, and even those who have published academic articles using these survey data, currently report very little about how their data were collected. Most commonly, researchers report some minimal information about how the sample was drawn and, roughly half the time, a response rate. At the time of writing, none of the major journals in political science, including the top journals in the field of public opinion, offer specific guidelines for reporting original survey research.

We join others (e.g., Hillygus 2009) in underscoring the need for scholars in the discipline to determine reporting standards for survey research. We believe that improved reporting would put a spotlight on the many important innovations being deployed by scholars around the world and help defend against skepticism regarding the quality of survey data from developing countries. Moreover, reporting standards would provide readers, reviewers, and data users with a common base of information, enabling them to assess academic survey data, utilize one another's techniques, and make productive methodological advancements.

We suggest as a starting point that published research using original survey data from the developing world—as well as the regional barometers and cross-national survey projects—ought to provide readers with the following information, where applicable:<sup>11</sup>

- Sample design protocol leading to the selection of individual respondents. This information should ideally include the specific strata used, sources for the sampling frame or population data, and the protocol employed for substituting enumeration areas, households, or individuals. It should also list eligibility criteria (e.g., citizens, adults, civilians) and exclusions (e.g., institutionalized individuals in hospitals, dormitories, or military bases).
- Complete questionnaire, including introduction, in the language(s) in which the survey was administered as well as the language in which the research is published.
- Translation protocol, including the language(s) of the survey, who translated the instrument, and how the translation was verified.
- Dates of survey fieldwork as well as dates and locations of pretests or pilots.
- Who supervised fieldwork—whether the researchers or a local survey firm or NGO—and how enumerators were trained.
- For face-to-face or phone surveys, how enumerators were selected.
- Platform and software used for enumeration.
- Procedures for validating or monitoring enumeration.
- Response rates and the method by which nonresponse information was collected.
- Whether post-stratification weights were used in the analysis and, if so, the weighting method, variables used, and sources of the population data.

Collecting survey data in developing countries is challenging, and researchers often innovate to gather high-quality data in these settings. These innovations, and more generally the methodological choices that inform how data are collected, represent important information for readers, reviewers, and data users. This information would not only allow the scholarly community to assess the appropriateness of the methods that were employed but also highlight the valuable innovations contributed by political scientists fielding surveys in the developing world and the high quality of the data that we believe most of their surveys are generating.

<sup>11</sup> Many researchers (including many of our expert respondents) hire local survey firms or NGOs to field their surveys and may not be involved in some aspects of data collection, such as sample design, enumerator selection, or enumerator monitoring. Still, these scholars would be able to request the relevant information from the implementing organization. Our proposed reporting standard could also be adopted for original survey data from the developed world, although some items (e.g., sample design, enumerator selection) would not apply to the opt-in online surveys that now predominate in those settings.



## LOOKING FORWARD

More and more political scientists are fielding surveys in the developing world. This growth has meant scholars are both confronting new challenges and developing solutions to them. Advances in satellite and remote sensing technology have helped researchers develop representative samples in the absence of reliable census data. Portable survey software and GPS have allowed them to monitor enumerators in real time and protect respondent information. Innovative solutions have improved the quality of survey data collected in environments where private interviews may be difficult to achieve, languages are diverse or oral, and political questions may be sensitive. Indeed, the high degree of creativity and resourcefulness of researchers conducting survey research in less-developed contexts is impressive.

Of course, there is still room for improvement. Many survey samples continue to rely on census data that are outdated and inaccurate. High-resolution satellite images are available commercially, but their cost far exceeds most survey research budgets. While remote sensing technology is developing rapidly, most survey researchers do not have the programming skills necessary to apply it to their needs. Moreover, there are still very few studies that actually validate these alternative sampling techniques. Increasing the accessibility of the technology to use satellite imagery to draw representative samples—and validating these methods—should be academic priorities.

Enumeration technology is also developing rapidly, and while it promises to improve the quality of enumerator-administered survey data, it can also be costly for individual researchers to develop and implement. Empirical evidence of the relative benefits associated with different oversight methods, from audio recording to GPS tracking, would help researchers make cost-effective choices that fit within modest research budgets.

Scholars have also been innovating with a wide range of strategies to address the challenges of administering surveys where, for a variety of reasons, respondents have difficulty responding to survey questions. It is crucial for researchers to think carefully about the population they are studying and ensure that their survey is as inclusive as possible. While many of these innovations seem likely to help mitigate the problem, we have little empirical evidence of their effectiveness and the magnitudes of their effects. Future methodological studies should compare various strategies and measure their effects.

We are particularly troubled by the finding that many academic survey respondents believe that the survey they are taking is sponsored by their government or another political actor. This is surely a finding that political scientists must take seriously. One of our expert respondents reported showing each survey participant an introductory video of the research team that emphasized the respondent's anonymity and the academic purpose of the study. Of course, this kind of effort may also emphasize the foreign sponsorship of a survey, which may itself bias responses (e.g., Cilliers et al. 2015). In developing regions where attitudes toward the United States or a former colonial power are especially salient, emphasizing that a survey is sponsored by a university based in one of those countries may introduce biases. But there is still little empirical evidence or guidance about the relative effects of these biases; this is one area where future research would be enormously useful.

We have also noted the absence of a common standard among political scientists working in the developing world for reporting the methods they employed in collecting their data. In an era notable for the simultaneous expansion of survey research, rapid pace of its methodological innovations, and skepticism regarding its quality, it is especially important that scholars clarify how their survey data were generated and highlight innovations. While we recognize that reporting requires additional efforts to describe a survey's design and implementation, we think the improved credibility that will come with better reporting more than offsets the cost. Moreover, researchers

interested in fielding their own surveys in developing contexts would benefit enormously from the ability to replicate—and improve on—the best practices of their colleagues.

## DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

## ACKNOWLEDGMENTS

We thank Liz Zechmeister for helpful comments on a draft of this article. For their feedback on our expert survey, we are grateful to Nara Pavão, Keith Weghorst, Yael Zeira, the colleagues who attended a workshop at Vanderbilt University, and the Working Group on Survey Methods in Developing Countries during the 2015 Annual Meeting of the American Political Science Association. We obtained human-subjects approval for the expert survey from the Institutional Review Boards at the University of Wisconsin–Madison and Vanderbilt University. Nicole Audette, Gui Russo, Rachel Schwartz, and Adam Wolsky provided excellent research assistance.

## LITERATURE CITED

- AAPOR. 2016. *Standard definitions: final dispositions of case codes and outcome rates for surveys*. Rep., 9th ed., Am. Assoc. Public Opin. Res., Oakbrook Terrace, IL. [http://www.aapor.org/AAPOR\\_Main/media/publications/Standard-Definitions20169theditionfinal.pdf](http://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf)
- Adida CL, Feree KE, Posner DN, Robinson AL. 2016. Who's asking? Interviewer coethnicity effects in African survey data. *Comp. Political Stud.* 49:1630–60
- Aronow PM, Coppock A, Crawford FW, Green DP. 2015. Combining list experiment and direct question estimates of sensitive behavior prevalence. *J. Surv. Stat. Methodol.* 3:43–66
- Banducci S, Stevens D. 2015. Surveys in context: how timing in the electoral cycle influences response propensity and satisficing. *Public Opin. Q.* 79:214–43
- Bauer JJ. 2014. Selection errors of random route samples. *Sociol. Methods Res.* 43:519–44
- Bauer JJ. 2016. Biases in random route surveys. *J. Surv. Stat. Methodol.* 4:263–87
- Beber B, Roessler P, Scacco A. 2014. Intergroup violence and political attitudes: evidence from a dividing Sudan. *J. Politics* 76:649–65
- Benstead LJ. 2014a. Does interviewer religious dress affect survey responses? Evidence from Morocco. *Politics Religion* 7:734–60
- Benstead LJ. 2014b. Effects of interviewer–respondent gender interaction on attitudes toward women and politics: findings from Morocco. *Int. J. Public Opin. Res.* 26:369–83
- Bischooping K, Schuman H. 1992. Pens and polls in Nicaragua: an analysis of the 1990 preelection surveys. *Am. J. Political Sci.* 36:331–50
- Blair G, Imai K. 2012. Statistical analysis of list experiments. *Political Anal.* 20:47–77
- Blaydes L, Gillum RM. 2013. Religiosity-of-interviewer effects: assessing the impact of veiled enumerators on survey response in Egypt. *Politics Religion* 6:459–82
- Bleck J, Michelitch K. 2017. Capturing the airwaves, capturing the nation? A field experiment on state-run media effects in the wake of a coup. *J. Politics* 79:873–89
- Blom M, Hox J, Koch A. 2007. The influence of interviewers' contact behavior on the contact and cooperation rate in face-to-face household surveys. *Int. J. Public Opin. Res.* 19:97–111
- Brislin RW. 1970. Back-translation for cross-cultural research. *J. Cross-Cult. Psychol.* 1:185–216
- Chauchard S. 2013. Using MP3 players in surveys: the impact of a low-tech self-administration mode on reporting of sensitive attitudes. *Public Opin. Q.* 77:220–31
- Chia SC. 2014. How authoritarian social contexts inform individuals' opinion perception and expression. *Int. J. Public Opin. Res.* 26:384–96

- Cilliers J, Dube O, Siddiqi B. 2015. The white-man effect: how foreigner presence affects behavior in experiments. *J. Econ. Behav. Organ.* 118:397–414
- Clayton A. 2015. Women’s political engagement under quota-mandated female representation: evidence from a randomized policy experiment. *Comp. Political Stud.* 48:333–69
- Corstange D. 2009. Sensitive questions, truthful answers? Modeling the list experiment with Listit. *Political Anal.* 17:45–63
- Davidov E, De Beuckelaer A. 2010. How harmful are survey translations? A test with Schwartz’s human values instrument. *Int. J. Public Opin. Res.* 22:485–510
- Davis DW, Silver BD. 2003. Stereotype threat and race of interviewer effects in a survey on political knowledge. *Am. J. Political Sci.* 47:33–45
- Diop A, Le KT, Traugott M. 2015. Third-party presence effect with propensity score matching. *J. Surv. Stat. Methodol.* 3:193–215
- Driscoll J, Lidow N. 2014. Representative surveys in insecure environments: a case study of Mogadishu, Somalia. *J. Surv. Stat. Methodol.* 2:78–95
- Druckman JN, Green DP, Kuklinski JH, Lupia A. 2006. The growth and development of experimental research in political science. *Am. Political Sci. Rev.* 100:627–35
- Durrant GB, Groves RM, Staetsky L, Steele F. 2010. Effects of interviewer attitudes and behaviors on refusal in household surveys. *Public Opin. Q.* 74:1–36
- Flores-Macías F, Lawson C. 2008. Effects of interviewer gender on survey responses: findings from a household survey in Mexico. *Int. J. Public Opin. Res.* 20:100–10
- Fujii LA. 2012. Research ethics 101: dilemmas and responsibilities. *PS Political Sci. Politics* 45:717–23
- Gelman A. 2016. Can you trust international surveys? *Monkey Cage Blog*, Feb. 27. <https://www.washingtonpost.com/news/monkey-cage/wp/2016/02/27/can-you-trust-international-surveys>
- Gingerich DW. 2010. Understanding off-the-books politics: conducting inference on the determinants of sensitive behavior with randomized response surveys. *Political Anal.* 18:349–80
- Gingerich DW, Oliveros V, Corbacho A, Ruiz-Vega M. 2016. When to protect? Using the crosswise model to integrate protected and direct responses in surveys of sensitive behavior. *Political Anal.* 24:132–56
- Glynn A. 2013. What can we learn with statistical truth serum? Design and analysis of the list experiment. *Public Opin. Q.* 77:159–72
- Gomila R, Littman R, Blair G, Paluck EL. 2017. The audio check: a method for improving data quality and detecting data fabrication. *Soc. Psychol. Pers. Sci.* 8:424–33
- Gonzalez-Ocantos E, de Jonge CK, Meléndez C, Osorio J, Nickerson DW. 2012. Vote buying and social desirability bias: experimental evidence from Nicaragua. *Am. J. Political Sci.* 56:202–17
- Groves RM, Couper MP. 1998. *Nonresponse in Household Interview Surveys*. New York: Wiley
- Groves RM, Couper MP, Presser S, Singer E, Tourangeau R, et al. 2006. Experiments in producing nonresponse bias. *Public Opin. Q.* 70:720–36
- Groves RM, Dillman DA, Eltinge JL, Little RJA, eds. 2002. *Survey Nonresponse*. New York: Wiley
- Groves RM, Fowler FJ Jr., Couper MP, Lepkowski JM, Singer E, Tourangeau R. 2009. *Survey Methodology*. Hoboken, NJ: Wiley. 2nd ed.
- Groves RM, Peytcheva E. 2008. The impact of nonresponse rates on nonresponse bias: a meta-analysis. *Public Opin. Q.* 72:167–89
- Harkness JA. 1999. In pursuit of quality: issues for cross-national survey research. *Int. J. Soc. Res. Methodol.* 2:125–40
- Harkness JA. 2007. Improving the comparability of translations. In *Measuring Attitudes Cross-Nationally: Lessons from the European Social Survey*, ed. R Jowell, C Roberts, R Fitzgerald, G Eva, pp. 79–93. Thousand Oaks, CA: Sage
- Harkness JA, Braun M, Edwards B, Johnson TP, Lyberg LE, et al., eds. 2010. *Survey Methods in Multicultural, Multinational, and Multiregional Contexts*. New York: Wiley
- Heath A, Fisher S, Smith S. 2005. The globalization of public opinion research. *Annu. Rev. Political Sci.* 8:297–333
- Heath A, Martin J, Spreckelsen T. 2009. Cross-national comparability of survey attitude measures. *Int. J. Public Opin. Res.* 21:293–315

- Hillygus DS. 2009. The need for survey reporting standards in political science. In *The Future of Political Science*, ed. G King, K Lehman Schlozman, N Nie, pp. 88–90. New York: Routledge
- Himelein K. 2016. Interviewer effects in subjective survey questions: evidence from Timor-Leste. *Int. J. Public Opin. Res.* 28:511–33
- Iannacchione VG. 2011. The changing role of address-based sampling in survey research. *Public Opin. Q.* 75:556–75
- Imai K, Park B, Greene KF. 2015. Using the predicted responses from list experiments as explanatory variables in regression models. *Political Anal.* 23:180–96
- Jann B, Jerke J, Krumpal I. 2012. Asking sensitive questions using the crosswise model: an experimental survey measuring plagiarism. *Public Opin. Q.* 76:32–49
- Johnson TP, Pennell BE, Stoop I, Dorer B, eds. 2018. *Advances in Comparative Survey Methods: Multicultural, Multinational and Multiregional Contexts*. New York: Wiley. In press
- Johnston R. 2006. Party identification: unmoved mover or sum of preferences? *Annu. Rev. Political Sci.* 9:329–51
- Kalton G. 1983. *Introduction to Survey Sampling*. Newbury Park, CA: Sage
- Kalton G, Flores-Cervantes I. 2003. Weighting methods. *J. Off. Stat.* 19:81–97
- Koczela S, Furlong C, McCarthy J, Mushtaq A. 2015. Curbstoning and beyond: confronting data fabrication in survey research. *Stat. J. LAOS* 31:413–22
- Krosnick J. 1991. Response strategies for coping with the cognitive demands of attitude measures in surveys. *Appl. Cogn. Psychol.* 5:213–36
- Kuriakose N, Robbins M. 2016. Don't get duped: fraud through duplication in public opinion surveys. *Stat. J. LAOS* 32:283–91
- Landry PF. 2010. Using clustered spatial data to study diffusion: the case of legal institutions in China. In *Contemporary Chinese Politics: New Sources, Methods, and Field Strategies*, ed. A Carlson, ME Gallagher, K Lieberthal, M Manion, pp. 219–45. New York: Cambridge Univ. Press
- Landry PF, Shen M. 2005. Reaching migrants in survey research: the use of the global positioning system to reduce coverage bias in China. *Political Anal.* 13:1–22
- Lee T, Perez EO. 2014. The persistent connection between language-of-interview and Latino political opinion. *Political Behav.* 36:401–25
- Legleye S, Charrance G, Razafindratsima N, Bohet A, Bajos N, Moreau C. 2013. Improving survey participation: cost effectiveness of callbacks to refusals and increased call attempts in a national telephone survey in France. *Public Opin. Q.* 77:666–95
- Liu M, Stainback K. 2013. Interviewer gender effects on survey responses to marriage-related questions. *Public Opin. Q.* 77:606–18
- Lohr SL. 2010. *Sampling: Design and Analysis*. Boston: Cengage Learning
- Lowes S, Nunn N, Robinson JA, Weigel J. 2015. Understanding ethnic identity in Africa: evidence from the Implicit Association Test (IAT). *Am. Econ. Rev.* 105:340–45
- Lupia A. 2015. *Uninformed: Why People Know So Little About Politics and What We Can Do About It*. New York: Oxford Univ. Press
- Lupu N, Peisakhin L. 2017. The legacy of political violence across generations. *Am. J. Political Sci.* 61:836–51
- Lyall J, Shiraito Y, Imai K. 2015. Coethnic bias and wartime informing. *J. Politics* 77:833–48
- Manion M. 2010. A survey of survey research on Chinese politics: What have we learned? In *Contemporary Chinese Politics: New Sources, Methods, and Field Strategies*, ed. A Carlson, ME Gallagher, K Lieberthal, M Manion, pp. 181–99. New York: Cambridge Univ. Press
- Massey DS, Tourangeau R. 2013. New challenges to social measurement. *Ann. Am. Acad. Political Soc. Sci.* 645:6–22
- McGorry SY. 2000. Measurement in a cross-cultural environment: survey translation issues. *Qual. Market Res.* 3:74–81
- Montalvo D, Seligson MA, Zechmeister EJ. 2018. Data collection in cross-national and international surveys: Latin America and the Caribbean. See Johnson et al. 2018. In press
- Mutz DC. 2011. *Population-Based Survey Experiments*. Princeton, NJ: Princeton Univ. Press
- Oliveros V. 2016. Making it personal: clientelism, favors, and the personalization of public administration in Argentina. *Comp. Politics* 48:373–91

- Olson K, Peytchev A. 2007. Effect of interviewer experience on interview pace and interviewer attitudes. *Public Opin. Q.* 71:273–86
- Paluck EL. 2009. Methods and ethics with research teams and NGOs: comparing experiences across the border of Rwanda and Democratic Republic of Congo. In *Surviving Field Research: Working in Violent and Difficult Situations*, ed. CL Sriram, JC King, JA Mertus, O Martin-Ortega, J Herman, pp. 38–56. New York: Routledge
- Peytchev A. 2009. Survey breakoff. *Public Opin. Q.* 73:74–97
- Robbins M. 2016. Yes, you can trust international surveys. Mostly. *Monkey Cage Blog*, Mar. 26. <https://www.washingtonpost.com/news/monkey-cage/wp/2016/03/26/yes-you-can-trust-international-surveys-mostly>
- Robbins M. 2018. New frontiers in detecting fabrication. See Johnson et al. 2018. In press
- Schaeffer NC, Dykema J, Maynard DW. 2010. Interviewers and interviewing. In *Handbook of Survey Research*, ed. PV Marsden, JD Wright, pp. 437–70. Bingley, UK: Emerald Group. 2nd ed.
- Seligson MA. 2008. Human subjects protection and large-N research: when exempt is non-exempt and research is non-research. *PS Political Sci. Politics* 41:477–82
- Simmons K, Mercer A, Schwarzer S, Kennedy C. 2016. Evaluating a new proposal for detecting data falsification in surveys: the underlying causes of “high matches” between survey respondents. *Stat. J. LAOS* 32:327–38
- Singer E, Couper MP. 2008. Do incentives exert undue influence on survey participation? Experimental evidence. *J. Empir. Res. Hum. Res. Ethics* 3:49–56
- Singer E, Hoewyk JV, Maher MP. 1998. Does the payment of incentives create expectation effects? *Public Opin. Q.* 62:152–64
- Sinibaldi J, Durrant GB, Kreuter F. 2013. Evaluating the measurement error of interviewer observed paradata. *Public Opin. Q.* 77:173–93
- Slomczynski KM, Powalko P, Krauze T. 2017. Non-unique records in international survey projects: the need for extending data quality control. *Surv. Res. Methods* 11:1–16
- Smith TW. 1997. The impact of the presence of others on a respondent’s answers to questions. *Int. J. Public Opin. Res.* 9:33–47
- Smith TW. 2009a. Codes of ethics and standards in survey research. In *The SAGE Handbook of Public Opinion Research*, ed. W Donsbach, MW Traugott, pp. 459–67. Thousand Oaks, CA: Sage
- Smith TW. 2009b. *A revised review of methods to estimate the status of cases with unknown eligibility*. Rep., Am. Assoc. Public Opin. Res., Oakbrook Terrace, IL. [https://www.aapor.org/AAPOR\\_Main/media/MainSiteFiles/FindingE.pdf](https://www.aapor.org/AAPOR_Main/media/MainSiteFiles/FindingE.pdf)
- Smith TW. 2010. Surveying across nations and cultures. In *Handbook of Survey Research*, ed. PV Marsden, JD Wright, pp. 733–63. Bingley, UK: Emerald Group. 2nd ed.
- Smith TW. 2015. Resources for conducting cross-national survey research. *Public Opin. Q.* 79:404–9
- SRC. 2016. *Guidelines for best practice in cross-cultural surveys*. Rep., 4th ed., Surv. Res. Cent., Inst. Soc. Res., Univ. Mich., Ann Arbor, MI. [http://ccsg.isr.umich.edu/images/PDFs/CCSG\\_Full\\_Guidelines\\_2016\\_Version.pdf](http://ccsg.isr.umich.edu/images/PDFs/CCSG_Full_Guidelines_2016_Version.pdf)
- Tourangeau R, Edwards B, Johnson TP, Wolter KM, Bates N, eds. 2014. *Hard-To-Survey Populations*. New York: Cambridge Univ. Press
- Tourangeau R, Smith TW. 1996. Asking sensitive questions: the impact of data collection mode, question format, and question context. *Public Opin. Q.* 60:275–304
- Tourangeau R, Yan T. 2007. Sensitive questions in surveys. *Psychol. Bull.* 133:859–83
- Tsai LL. 2010. Quantitative research and issues of political sensitivity in rural China. In *Contemporary Chinese Politics: New Sources, Methods, and Field Strategies*, ed. A Carlson, ME Gallagher, K Lieberthal, M Manion, pp. 246–65. New York: Cambridge Univ. Press
- West BT, Blom AG. 2017. Explaining interviewer effects: a research synthesis. *J. Surv. Stat. Methodol.* 5:175–211
- West BT, Kreuter F. 2013. Factors affecting the accuracy of interviewer observations: evidence from the National Survey of Family Growth. *Public Opin. Q.* 77:522–48

- West BT, Olson K. 2010. How much of interviewer variance is really nonresponse error variance? *Public Opin. Q.* 74:1004–26
- Zechmeister EJ. 2016. Ethics and research in political science: the responsibilities of the researcher and the profession. In *Ethics and Experiments: Problems and Solutions for Social Scientists and Policy Professionals*, ed. S Desposato, pp. 255–61. New York: Routledge
- Zechmeister EJ, Seligson MA. 2012. Public opinion research in Latin America. In *Routledge Handbook of Latin American Politics*, ed. P Kingstone, DJ Yashar, pp. 467–82. New York: Routledge





# Contents

Politics and Political Science <i>Robert Jervis</i> .....	1
A Conversation with Charles V. Hamilton <i>Charles V. Hamilton and Fredrick C. Harris</i> .....	21
Transparent Social Inquiry: Implications for Political Science <i>Colin Elman, Diana Kapiszewski, and Arthur Lupia</i> .....	29
Political Trust in a Cynical Age <i>Jack Citrin and Laura Stoker</i> .....	49
State Capacity Redux: Integrating Classical and Experimental Contributions to an Enduring Debate <i>Elissa Berwick and Fotini Christia</i> .....	71
Unwelcome Change: Coming to Terms with Democratic Backsliding <i>David Waldner and Ellen Lust</i> .....	93
Cities and Politics in the Developing World <i>Alison E. Post</i> .....	115
A Taxonomy of Protest Voting <i>R. Michael Alvarez, D. Roderick Kiewiet, and Lucas Núñez</i> .....	135
Confucian Political Theory in Contemporary China <i>Yi-Huah Jiang</i> .....	155
On the Theory of Parties <i>Nolan McCarty and Eric Schickler</i> .....	175
Advances in Survey Methods for the Developing World <i>Noam Lupu and Kristin Michelitch</i> .....	195
Complicit States and the Governing Strategy of Privilege Violence: When Weakness Is Not the Problem <i>Rachel Kleinfeld and Elena Barham</i> .....	215
How to Think About Social Identity <i>Michael Kalin and Nicholas Sambanis</i> .....	239
Some Advances in the Design of Survey Experiments <i>Paul M. Sniderman</i> .....	259

The Other Side of Taxation: Extraction and Social Institutions in the Developing World <i>Ellen Lust and Lise Rakner</i> .....	277
Redistribution Without a Median Voter: Models of Multidimensional Politics <i>Torben Iversen and Max Goplerud</i> .....	295
Political Psychology in International Relations: Beyond the Paradigms <i>Joshua D. Kertzer and Dustin Tingley</i> .....	319
Violent Conflict and Political Development Over the Long Run: China Versus Europe <i>Mark Dincecco and Yubua Wang</i> .....	341
Does Foreign Aid Build Peace? <i>Michael G. Findley</i> .....	359
Political Theories of Migration <i>Sarah Song</i> .....	385
Legitimacy in Areas of Limited Statehood <i>Thomas Risse and Eric Stollenwerk</i> .....	403
Dead But Not Gone: Contemporary Legacies of Communism, Imperialism, and Authoritarianism <i>Alberto Simpser, Dan Slater, and Jason Wittenberg</i> .....	419
Models of Other-Regarding Preferences, Inequality, and Redistribution <i>Matthew Dimick, David Rueda, and Daniel Stegmueller</i> .....	441
Radicalization: A Relational Perspective <i>Donatella della Porta</i> .....	461
Justice and Future Generations <i>Simon Caney</i> .....	475
Theories of Institutional Corruption <i>Dennis F. Thompson</i> .....	495
International Negotiation: Some Conceptual Developments <i>Barry O'Neill</i> .....	515

## Errata

An online log of corrections to *Annual Review of Political Science* articles may be found at <http://www.annualreviews.org/errata/polisci>