

# MODULE V

## Developing and Using Population Surveys



**PURPOSE:** There are exciting new innovations in population-based surveys to assess change in FCV settings. Such tools can provide critical insight on the livelihoods and perceptions of a wide range of beneficiary (and non-beneficiary) groups. Part of the hesitancy to use these tools is because their purpose and methodology is not well understood. The Guidance fills this knowledge gap by reviewing how population surveys are designed, the specificities of sampling, the preparation of questionnaires, and the interpretation of results. It also lays out some limitations of these approaches, questions related to ethics and some insights into emerging techniques.

Population surveys are increasingly used to evaluate programs and monitor broader economic, social, and institutional dynamics in FCV contexts.<sup>32</sup> This module of the Toolkit provides an overview of the applications of population surveys. It presents lessons on survey design and implementation, with a focus on managing the particular challenges introduced by conflict, violence, and institutional weakness. It also emphasizes the need for teams to identify potential sources of bias at each stage of survey design and implementation, both in order to generate high-quality data and to use as a guide in interpreting and using survey data from FCV environments.

---

32 A population survey uses a random sample of interviews in order to estimate beliefs, experiences, perceptions, and attitudes in the population. Surveys are sometimes referred to as “perception surveys,” particularly when the focus is on tracking population attitudes and beliefs; but even “perception” surveys typically collect much more than perception data, such as information on each respondents’ location, socioeconomic status, and experiences.

## Why Use Survey Data to Monitor Progress on Peacebuilding and State-Building?

Surveys have long been used in middle- and high-income countries to gauge public mood and evaluate and communicate the results of policies and reform processes. There is growing appetite for the application of surveys to track key metrics of quality of life, satisfaction and attitudes, and confidence in transition processes in countries affected by FCV. Surveys can be used to gauge public experience and to assess progress toward policy goals. Accordingly, small and large-scale household surveys have been used in many FCV settings—from active conflict areas to post-conflict states—for a wide range of purposes.

First, population surveys have been used in low-income and fragile countries to **generate data on development trends** and outcomes. As noted in Module IV, data in FCV contexts are often fragmented, limited, or low-quality. Survey data can be used to fill-in information gaps that would be provided by more robust and routine administrative statistics. Survey data can also complement other forms of monitoring data. For example, the quality of service delivery can be examined by linking data on health or education facilities with perception surveys of end users. Triangulation using survey and administrative statistics can improve the reliability of analyses, and can enrich evaluations by helping to unpack and explain associations observed in other quantitative data.

A second and arguably more salient use of survey data in FCV contexts is to **provide insights into citizens' lives, experiences, and perceptions**. In FCV contexts, institutional frameworks for managing conflict are often weak, so citizens' perceptions of the state's legitimacy and confidence in reform efforts can be especially important. More generally, survey data can complement administrative indicators by providing insights into how institutions are perceived and experienced by citizens. For example, surveys can provide data on experiences of public sector corruption; citizens' perceptions of police forces; and access to justice.

Survey data can also be **disaggregated**, broken down to identify differences in experiences and attitudes across gender, age, and other identity groups. Disaggregation can help inform policymakers, by showing them whether the state is perceived as more or less legitimate or effective by different groups. Surveys can also be used to track population priorities in transition environments when leadership often must choose between an overwhelming number of urgent and competing tasks.

Surveys also are sometimes used to **focus attention on important or vulnerable groups** that might not be effectively tracked by either conventional administrative statistics or other forms of monitoring data. Some of these populations may be hidden yet pivotal, such as former combatants, whose economic and



### Box 9. Using Surveys to Measure Strategic and Operational Indicators

Surveys can generate data useful for tracking trends at both strategic and operational levels. As with other indicators, survey data should be designed with context in mind, and should be informed by specific goals, targets, and theories of change. For example, a program may be designed to provide services to vulnerable ethnic minority populations, to support a nascent peace process.

Program	Operational indicators	Objective indicators	Strategic indicators
<i>Peace process support: service delivery in ethnic minority areas</i>	Percent of ethnic minority households reporting that eligible children are attending primary school (outcome)	Change in number of primary schools in ethnic minority provinces (output)	Percent of ethnic minority members who say that they support the peace process
	Percent of ethnic minority members reporting access to improved water source (outcome)	Number of wells constructed (output)	Percent of ethnic minority members reporting confidence in the government

At the operational level, survey data might provide insights into changes in access to services and service quality: whether ethnic minority households report that children are able to attend primary school, or whether more ethnic minority respondents report having access to improved water and sanitation systems, compared to before the program launched. The program might also have impacts on broader strategic issues, such as levels of confidence in the peace process among minority populations, or perceptions among minority group members that the government is invested in their wellbeing.

social reintegration can be central to peace and stability. Finally, surveys are frequently used to evaluate the impact of a particular project or intervention. The World Bank and partner organizations have invested significant resources in survey-based evaluation, notably in the case of large-scale community-driven development programs. When used for evaluation, multiple rounds of surveys are typically deployed: an initial survey to establish a baseline measure, with follow-up surveys tracking change.

## Deciding Who—and How—to Survey

A key advantage of population surveys is that they use a relatively limited form of data collection—a set of interviews conducted with a random sample of people—to draw inferences about much larger groups.<sup>33</sup> It is critical to define clearly the **population**, the larger group on which information is sought.<sup>34</sup> There is no preset definition for a population; it depends entirely on the goals of the survey. A population can be the entire population of a country, a geographic area (such as disaster-struck provinces in Pakistan), a social group of interest (such as former child soldiers in Uganda), or people receiving a program intervention. In short, the population is defined by the group or intervention under consideration.

Once the population has been identified, a sampling design can be developed. **Sampling** is the process of selecting units—such as households, villages, or individuals—from the population to interview. A key step in the design process is to assemble the **sampling frame**, the list from which the sample will be randomly selected. A lot hinges on the quality of the sampling design: a frame with missing or inaccurate elements will yield a flawed sample and an inaccurate picture of the population.

For this reason, it is advisable to consult with a sampling statistician, and to ensure that the design is tailored to the context and needs of the study.<sup>35</sup> This also requires those who undertake the study to be very clear on the policy and programming outcomes the study is supposed to inform. In addition to clearly specifying the overall population of interest, it is important to identify any relevant subpopulations: particular groups or areas across which needs, experiences, or other factors salient to program design, implementation, or evaluation may matter.

Sampling designs usually reflect a difficult balance between the goals of the survey, implementation considerations, and budget constraints. Several sampling approaches are commonly used in FCV environments:

- 
- 33 The exception is a census, which is a full enumeration of all people in a given country or locality. Population censuses are expensive, logistically challenging, and time-consuming, which is why they are typically done only once a decade in middle- and high-income contexts. In FCV, censuses occur with lower frequency, and can have missing data owing to lack of physical access, or other data structure and quality problems.
  - 34 This guideline may sound basic, but it is sometimes ignored. For example, a national survey might intend to cover the entire population of a given country; but some subpopulations, like nomadic groups or displaced people, may not be covered in a household survey. A household survey of young men might miss men who are in prison or in the military. A clear definition can help identify who is in—and what, if anything, is necessary to ensure proper coverage.
  - 35 It is also important to document the sampling design and selection process carefully so that sampling weights and the margin of error can be calculated.



- In **stratified random sampling**, the population is divided into a set of strata (subgroups), which are then each independently sampled. This approach can ensure that sample sizes are appropriate for important subgroups in the population, and that smaller groups (for instance, ethnic minorities) receive adequate coverage.
- In multistage **cluster sampling**, random samples are drawn from a set of nested geographic areas, such as a set of provinces; the districts within each province; the villages within each district. Cluster sampling is often used when there is no adequate sampling frame; it can also help control costs, since travel time within clustered points is generally reduced. That said, clustered areas—say, villages—are also typically more homogenous than non-clustered areas; care should be taken to balance cluster number and diversity.

The population and sampling design define the spatial dimensions of the survey. But it is important to think about the temporal dimension: *whether and how to measure over time*. Most population surveys are **cross-sectional**: they capture a slice of a population's opinion and experience in a particular place and time. Taking **repeated cross-sections** can help measure change over time. **Longitudinal (or "panel")** surveys interview the same people (or households) repeatedly, through multiple survey "waves." Longitudinal studies can help track micro-level changes, and can provide sharper insights into the effect of external factors such as shocks or social programs on the lives of individuals. They can also help measure programs' efficacy over time, or look for second- and third-order effects.

However, longitudinal surveys are typically more expensive and difficult to implement, since the same interview subjects must be tracked down each time that the survey is run. Attrition—people dropping out of the survey because they cannot be located, die, stop wanting to participate—is a problem in middle- and high-income contexts, but can be especially challenging in FCV contexts, where population mobility and mortality are often much higher. Security concerns, including the need to retain data to identify specific people to take part in the survey, can also pose unique risks in FCV environments. Despite the challenges, panel surveys are underway in multiple FCV contexts. New technologies, such as low-cost GPS devices to record interview sites, and the diffusion of mobile phones, can make it easier to locate respondents for follow-up interviews (see Module VI).

## Challenges in FCV Contexts

Fragile and conflict-affected areas are challenging environments for survey work.<sup>36</sup> The two main challenges encountered in FCV are lack of data and logistical and access constraints created by insecurity. Dealing with these issues requires planning and risk management, but despite these challenges, ambitious surveys can be fielded in very difficult environments.

Data limitations are primarily a challenge for sampling design. High-income countries usually have a wealth of data to use in building a sampling frame, such as well-organized lists of administrative and geographic areas, with up-to-date socio-demographic information. **Data necessary for sampling are often unavailable in FCV areas**, due to a lack of recent or reliable census information, or fragmentation of records. Various approaches, ranging from cluster sampling to the use of remote sensing tools to build a roster of households can provide a basis for rigorous sampling design even in the absence of administrative data.

Insecurity in FCV contexts can pose significant challenges, for both data collection and data integrity. Access constraints are the most visible problem. **Active conflict or hostile non-state armed groups can make some areas inaccessible.** This risk is not always evenly distributed: even within a single conflict-affected district, some areas may be safe to conduct fieldwork, while others may be too insecure. When an area is inaccessible due to security conditions, it is often replaced with the nearest secure sampling point—typically one that is as similar as possible.

The challenge is that replacement can introduce bias. The population in an insecure area may differ in important ways from a similar replacement.<sup>37</sup> While statistical procedures can help address such bias, it is difficult to eliminate. It may be more effective and useful to consider the likely *direction* of the bias, to interpret the results with care, and to be cautious regarding generalization to the original population. In short, always ask, what is this sample actually representative of?

---

36 Some challenges, such as illiteracy and innumeracy, are common to both FCV environments and other low-income (or rural) contexts. These topics are treated elsewhere, so this section focuses on challenges that are distinctive to FCV.

37 Patterns of conflict can also introduce more subtle forms of bias, which can be difficult to detect and threaten the integrity of survey data in FCV contexts. A recent analysis of mortality studies in Iraq found evidence that a sampling technique—selecting a random main street, and subsequently a random cross street on which to begin selecting households for interviews—generated biased estimates of violence. Why? Police patrols, roadblocks, and other factors likely to attract violence also congregated along the intersection of major roads and side streets, while areas away from major roads were much less affected by violence. Understanding variation in conflict can help configure sampling designs down to the last, often crucial, steps. See Johnson et al. (2008).



## Box 10. Undertaking Surveys in Challenging Environments

Some surveys have taken creative approaches to dealing with data gaps and operational insecurity in FCV environments. A recent survey of Mogadishu applied a novel approach to constructing a sample frame, so that a rigorous sample could be drawn even without census data or an enumeration of buildings within the city. This approach used satellite imagery to construct a map of Mogadishu and estimated population size based on each dwelling to construct a sample frame (see Module VII).

Creative methods have also been applied to finding “hidden populations.” A survey of former combatants in Sierra Leone faced a sampling challenge: although a disarmament, demobilization, and reintegration (DDR) program had taken place, there was no complete list of former combatants. The survey team sampled a random set of urban and rural locations, and asked knowledgeable people from the area—local chiefs and DDR program staff—to identify former combatants living in the locality.

Operational security in FCV environments can pose major challenges to the safety of the survey team. A recent Asia Foundation survey in the conflict-affected deep south of Southern Thailand managed insecurity by linking a national survey firm with local NGOs from the survey area. NGO staff accompanied enumerators to villages within active conflict areas, introducing the survey team to village officials and influential persons, and helping to build trust and broker access. Careful enumerator recruitment can also help build security: an upcoming World Bank survey in Mogadishu, Somalia, has recruited enumerators from across the city, and from a range of locally salient clans; this approach helps to ensure that enumerators are known, trusted, and can access otherwise insecure areas to safely gather an unbiased sample of respondents.

Even when an FCV area is *relatively secure*, **negotiating access to implement a population survey can be challenging**. Survey teams frequently have to overcome mistrust among local populations, who may be concerned about the presence of strangers asking questions. Working through local intermediaries, such as NGOs and civil society organizations, may help secure access. Particularly in rural areas, it may also be necessary to gain the formal consent of local leaders in order to conduct interviews. However, it is important to note that seeking formal consent also means announcing the presence of the survey—which can potentially compromise the anonymity of the respondents.<sup>38</sup>

---

38 Even if access is obtained, FCV environments may induce biased responses. Mistrust of outsiders can lead civilians to withhold or shape their responses in inaccurate ways. Potential or perceived threats from actors with political, social, economic, or military strength can similarly shape responses.

## Lessons on the Design Process

Principles for survey design in FCV contexts are in many ways similar to good practices in stable low- and middle-income areas. But there are additional challenges, including (typically) limited experience and institutional knowledge on survey deployment in FCV localities, and limited local capacity and training in survey methodologies and implementation. Additional preparation and care in the design process can go a long way toward addressing these challenges.

Some general strategies to keep in mind:

- **Start early** and ensure that adequate resources and time have been allotted so that the survey can go through an **iterative design process**. Surveys perform best when the sampling design, questionnaire topics, and indicators are refined through multiple rounds of feedback and testing. Almost invariably, surveys in FCV contexts require flexibility and adaptation.
- **Clearly define the research question** or problem. What is the survey intended to explain or explore? Researchers face the (understandable) temptation to add a host of questions to any survey in an effort to gain as much information as possible. A sharp focal question can help identify the critically important topics, questions, and outcome measures. Without a clear focal question, key topics and measures can be missed, or a survey can become unmanageably long.
- **Invest in exploratory research**. Qualitative tools, such as semi-structured interviews,<sup>39</sup> can be useful both in identifying the key questions to ask, how to ask them so that they are understood by a wide range of people, and in developing appropriate response options. Early qualitative research can also ensure that questions reflect issues relevant to women, minorities, and other social groups of concern.
- Iteratively **pilot and refine the questionnaire**. It is critically important to test and refine the survey before it goes to the field. Pilot tests consist of a small number of interviews conducted in the field, preferably in areas where the survey will eventually be fielded. They can also be combined with in-depth focus groups that discuss the design and comprehensibility of survey questions.<sup>40</sup> A pilot can provide important early indications of how the survey will

---

39 Semi-structured interviews can be broadly defined as conversations that have a set of guiding questions or themes, but are open to new topics. There are several relevant types. **Key informant interviews** are in-depth, one-on-one conversations, often with respondents who will have significant knowledge on a given topic. Focus groups are larger conversations in which an interviewer moderates and guides a group discussion. **Focus groups** can be designed to explore the opinions of a particular demographic (e.g., male youths, female entrepreneurs) or can be mixed to identify variation in opinions or experiences across multiple groups.

40 The exact number of pilot interviews depends upon the goals of the survey and whether a ro-



perform in the field, and whether respondents understand questions. Very often some question designs—wordings, structure, response options—will need to be revised based on the pilot.

- **Select enumerators carefully.** The ideal team of interviewers will be educated, familiar with the area where the survey will be conducted,<sup>41</sup> conversant in all local languages, and will have a mix of men, women, and relevant social groups. In practice, this ideal is very difficult to meet. Tradeoffs are inevitable. In insecure environments, local familiarity and contacts may be the most important factor. Depending on cultural context, it may be critical to ensure that female respondents can be interviewed by female enumerators (or that members of a particular ethnic group are interviewed by co-ethnics). Above all, team composition should be driven by preparatory fieldwork.
- Pilots also offer an opportunity to **train enumerators and field supervisors**, to ensure that they select households and respondents correctly, and are familiar and comfortable with the survey instrument. At minimum, enumerators should be trained through practice interviews, and should have a clear understanding of the sampling design and the importance of randomized selection.<sup>42</sup>
- **Enumerators should also be trained to answer questions from local actors**, powerbrokers, and respondents that inevitably arise once they are in the field. Responses should be selected to reduce possible biases, protect survey respondents and personnel, and to calibrate expectations appropriately regarding the potential for survey data collection to yield a development project.

---

bust sample of subgroups (youth, women, minority groups) is desired. Generally, pilots should cover at least 25 people; larger numbers can provide more detail on how the questionnaire will eventually perform, and can allow disaggregation to subgroups, but are more costly.

- 41 Survey experts often recommend that interviewers be recruited from the area where the survey will take place. But there is a delicate balance: enumerators should be familiar enough with the area that they can operate securely and effectively, but should still be strangers, so that respondents feel comfortable disclosing sometimes-sensitive information (on income, perceptions, and the like.)
- 42 The practice of randomization is often strange to non-statisticians, but enumerators need to understand its importance, particularly in the face of reasonable-seeming alternatives they could implement once in the field, such as selecting a closely grouped set of houses for interviews, or interviewing the first person to answer the door within a household.

## Developing the Questionnaire

Questionnaire design should be carefully tailored to the local context, and should be informed by relevant social, institutional, and political economy analyses. However, this does not mean that every survey should be developed from a blank slate. It is advisable to identify potential sources of question designs that have already been vetted, and which may be adapted. This can include prior surveys from the same context, surveys deployed in similar environments, or questionnaires addressing similar issues. Almost invariably, new question designs will be necessary; these should be extensively tested and revised. Design can also draw from previous data collection efforts that have proven flawed or shown quality concerns in the data collection.

Whether adapted or bespoke, there is a wide range of potential question designs that can be used to gather data. Some widely employed examples include:

- **Perception** (or attitudinal) questions are used to assess how people *think* and *feel* about an issue, institution, or situation. They are typically used to gather information on subjective issues, such as trust, satisfaction, perceived efficacy, and legitimacy. When used in combination with objective data sources, perception questions can yield insights into why people feel as they do. However, perceptions can change rapidly, so perception data should be considered in both temporal and spatial context.
- **Behavioral** (or experiential) questions are used to map and understand *how people act*: whether they vote, interact with members of other identity groups, how often they interact with police and justice institutions. They can be especially helpful in tracking changes in concrete actions over time, and in assessing the impact of interventions. When designing behavioral questions, it is important to bear in mind that answers are *self-reported behaviors*, which are subject to social desirability bias and faulty memory.<sup>43</sup>
- **Incident reporting** questions ask a respondent to report on the incidence or frequency of distinct events within a span of time. Incident reporting frequently taps additional information, including the intensity, cause, and consequences of events. Incident reporting questions are commonly used in surveys of crime and victimization.

---

43 The passage of time can cause respondents to forget details of the past: when an event happened, what security conditions were like, what livelihood opportunities were available at a given point in time. Recall bias can be random, but there is concern that where respondents have forgotten details, they may “construct false memories” using contemporary narratives or media accounts. Stress and trauma can also lead to bias. See Schulhofer-Whol (2014) and Bradburn et al. (1987).



### Box 11. Perception Surveys—Afrobarometer

There is frequently a gap in objective and subjective data on the quality, quantity, and availability of services in FCV settings. Perception-based data on attitudes and behavior can provide important insights into the overall experience of citizens, a key priority identified in the 2011 WDR. There is a tremendous demand from governments, donors, civil societies, and the media for reliable data on citizen perspectives. It can provide insights into the extent of progress, but also shortcomings.

The Afrobarometer is one of a number of networks collecting comparative data on public attitudes through routine surveys. It has administered four rounds of surveys, with its last series of assessments (2008–2011) covering 5 of the 17 FCS in Sub-Saharan Africa (Burkina Faso, Liberia, Madagascar, Mali, and Zimbabwe). FCS settings have traditionally been excluded owing to the additional logistical burden and constraints associated with surveys. However, with support from the World Bank SPF, the Afrobarometer coverage is expected to expand considerably from 2012–2015, including to Burundi, Côte d’Ivoire, Mali, Niger, Sierra Leone, South Sudan, and Togo.

The Afrobarometer collects data through large-scale household surveys. Through its existing network, it undertakes survey design, pretests and translation, training of fieldworkers, sampling design, survey administration, data collection, and the rest. Since its creation, it has fielded more than 70 surveys across 35 countries. Key focus themes include democracy, governance, livelihoods, conflict and crime, levels of participation, identity questions, taxation, gender, and social service delivery. Some indicators used in the past relate to how safe people feel, experiences with crime and violence, and perceptions of the role and effectiveness of security institutions.

Lastly, surveys should include **socio-demographic** questions to collect basic data on the respondent or household’s characteristics, such as age, gender, employment, income, education, social identity.<sup>44</sup> These indicators provide dimensions along which survey data can be disaggregated, as well as statistical control variables for quantitative analysis.

---

44 Social identity is often considered a demographic topic, but frequently involves complex dimensions of perception and may need to be measured through a range of techniques.

While question design is a continuously evolving art, number of lessons are broadly accepted:

- **Triangulate.** It can be helpful to use a variety of question types to measure critical topics. A mix of behavioral, perceptual, and experiential questions can capture multiple dimensions of a complex issue. For example, public-sector corruption could be measured along multiple axes: perceptions of its intensity, population-level exposure, willingness to pay a bribe under given conditions, among others.
- **Avoid jargon.** Survey questions should be simple and understandable to all potential respondents. Avoid acronyms or technical terms. Consider how respondents with limited literacy or numeracy will interpret questions.
- **Be specific.** Vague questions can yield unreliable answers, since different people may interpret a question in very different terms. For example the simple perception question, “To what extent do you trust the police?” leaves much room for interpretation. Do you trust the police to investigate a problem fairly? Efficiently? Promptly? Without bias? These dimensions may correlate—or may not.
- **Avoid double-barreled questions.** Double-barreled questions contain two distinct topics. For example, “Do the courts process criminal cases fairly and efficiently?” Here, the word “and” is a useful signal that a question is measuring multiple concepts, and should be broken into two items.
- **Avoid leading questions or framing** that might bias a response. Survey questions can often contain assumptions that are not shared by all participants. For instance, asking “in order to expand government primary education, should we raise taxes?” assumes that the government is the preferred service provider.
- **Make the response format obvious.** It is not always obvious how a respondent should answer a question. “How far are you from the nearest primary school?” could be answered in linear distance, or travel time; “What is your income?” could be specified temporally (last week, last month, annually), and by scope (personal, household).
- When asking agree/disagree questions **include “don’t know” and mid-point options.** Some respondents may simply not be familiar with an issue, or may not have an opinion at all, while others may be agnostic. These are different states, but may be recorded identically unless both response options are clearly available. Be aware, however, that “don’t know” responses may be disproportionately chosen in response to sensitive questions—not because a respondent does not know, but rather to avoid providing information on a sensitive issue.
- **All coding options should include a “refuse” response.** Interviewers should be able to indicate when a respondent simply refuses to answer a



question.<sup>45</sup> This can be critical in identifying sensitive topics.

- **Use open-ended questions strategically.** Most questions are closed: they have a defined and limited set of response options. Open-ended questions allow a participant to respond in her own words; this can yield valuable information, especially when the range of potential responses is unknown, but adds a significant burden when coding and analyzing responses. Open-ended questions can also be used strategically in survey design—even if their answer is not analytically useful, they can provide an entry point into more sensitive questions.

A few final words on questionnaire design are in order. First, **translate with care.** The tone or even basic concepts in a question can be warped in translation. “Back-translating” the questionnaire from the target language to original language can help identify points of slippage; a detailed conceptual and tonal review of each question by the native speakers on the survey team is also strongly advised. Second, **use feedback from enumerators.** Field interviewers can provide valuable insights into how questions are being processed and understood by respondents, which can be vital in revising question designs. Third, **limit survey duration.** It can be tempting to capture as much information as possible in a single survey. However, answering structured questions takes energy, especially if specific details about past events and experiences need to be recalled. People can become mentally exhausted during a long interview, leading to random or incorrect answers.

## Dealing with Sensitive Issues

Many of the most important issues for monitoring in FCV contexts may also be among the most sensitive: state legitimacy, discrimination, social cohesion, violence. Even relatively straightforward socio-demographic questions, such as a person’s age or income can be surprisingly sensitive. While there is no simple answer to how well surveys can measure sensitive topics, an emerging body of evidence suggests that direct questions can be problematic.

It is important to keep in mind that surveys are usually conversations between strangers. Without a baseline level of trust and familiarity, respondents can be understandably hesitant to reveal views and experiences on stressful or intimate topics. Social desirability bias—a psychological tendency to avoid reporting embarrassing or unpleasant beliefs, or to over-report positive ones—can lead people to refuse to answer sensitive questions, or to give inaccurate information.<sup>46</sup> Both nonresponse and inaccurate reporting can systematically bias survey results, sometimes significantly. In FCV settings, disclosing answers to

---

45 This option should, of course, not be read, but simply marked as needed by the enumerator.

46 See Parks et al. (2013).

survey questions may also be considered dangerous for respondents, depending on political, security, and social dynamics.

Which issues are sensitive? It depends upon context and culture. In areas that have experienced ethnic conflict, questions about social trust can be problematic. Population displacement can make ordinarily unproblematic questions about land tenure and ownership sensitive. Some sensitivities can be predicted, but many cannot. So, in FCV contexts, careful preparatory work—such as interviews and focus groups—should take place early on in the design process. If possible, Bank staff responsible for survey design should participate or observe, and watch body language, expression, and tone. It can be especially helpful to have local experts and counterparts take part in this iterative testing process, since their greater familiarity with local norms, customs, and social signals can help identify more subtle problems, help identify unexpectedly sensitive questions, and aid in rephrasing them.

Sensitive topics generally fall into two categories: issues that are taboo or uncomfortable to discuss, and issues about which the respondent would fear disclosure of their opinions or experiences.<sup>47</sup> While there is no perfect solution to these challenges, there are techniques and tools to address each type of sensitivity. Following are a few useful strategies.

**Try to ensure privacy.** Surveys are usually designed to be one-on-one conversations between the interviewer and respondent; without privacy, most people are not as candid or open. Privacy is also necessary to protect respondents. In many low-income contexts, privacy can be difficult to achieve, as family members, neighbors, or curious members of the community may try to observe the conversation. In FCV environments, local actors or authorities can be sensitive about strangers collecting data, and might try to monitor who is being interviewed, or the interviews themselves. Some strategies can help: before entering a survey site, have field supervisors meet with local actors or authorities to explain the goals of the survey, and the importance of privacy. Have enumerators pause the interview if another person interrupts. It can also be helpful to have multiple interviewer teams run parallel interviews at a survey site, so that individual teams are harder to track or monitor.

**When asking direct questions, it can help to acknowledge the sensitivities up front.** For taboo or embarrassing issues, question wording can help mitigate social desirability bias, by signaling that multiple answers are valid and acceptable, and that the question is not normative. This might seem self-evident, but questions often implicitly signal that there is a right answer. For example, a

---

47 Some issues—such as domestic violence—may fall into both categories. Roger Tourangeau and Ting Yan “Sensitive Questions in Surveys” *Psychological Bulletin*, Vol. 133, No. 5, 2007. Accessed at: [http://www.learnlab.org/research/wiki/images/a/a8/Tourangeau\\_SensitiveQuestions.pdf](http://www.learnlab.org/research/wiki/images/a/a8/Tourangeau_SensitiveQuestions.pdf)



simple polling question on political participation—“Did you vote in the last election?”—can yield inaccurate data, because norms suggest that good citizens ought to vote. Adjusting question wording to acknowledge that there are valid reasons for not voting can help mitigate bias. For example, “Sometimes people are unable to vote, because of registration problems, distance to the polling place, or other factors. Were you able to vote in this election?”<sup>48</sup> This is equally important for factual questions: respondents rarely want to look ignorant in front of an interviewer, so acknowledging that it is normal or reasonable not to know an answer can help reduce nonresponse or random answering.

**Self-administered items**—questions or short modules that respondents fill out, rather than being asked by an interviewer—**can help address taboo or embarrassing topics**. In pen and paper surveys, self-administered questions are sometimes sealed in an envelope, emphasizing that the information will remain private.<sup>49</sup> Some evidence from research in high-income countries suggests that computer-aided self-administered questions—for example, administered by smartphone—can perform better.

**Indirect questions**, also sometimes described as “survey experiments” can help get traction on both taboo and potentially dangerous topics. Indirect question designs do not require that a participant give her individual opinion or experience on a sensitive issue. Several of the most common approaches, including the “double list” technique, use randomization to estimate the *overall percentage* of people who have a given view. In theory, respondents answer truthfully, since there is no risk of individual disclosure.

Such questions have been used to estimate support for government and insurgent forces in active conflict areas, and attitudes toward transitional justice in post-conflict zones. However, it is worth noting that indirect questions can be difficult to implement in practice. A larger sample may be necessary to boost statistical power. Experiments typically require extensive pretesting and training for interviewers, and often need to be clearly explained to respondents. Finally, techniques such as the double list are useful when population proportions—such as the percentage of people who have paid a bribe—are of interest, but do not shed light on individual-level beliefs or experiences.

---

48 See Duke Initiative on Survey Methodology “Tipsheet: Sensitive Questions.” Retrieved at: <http://www.dism.ssri.duke.edu/pdfs/Tipsheet - Sensitive Questions.pdf>

49 In this approach, the self-administered items and the remainder of the survey data are identified by a unique identification number, so that they can later be linked for analysis.

## New Methods

The rapid rise of information and communication technologies (ICT) has opened up new avenues for survey research in fragile contexts. While ICT-enabled tools are not a complete solution, they can help to mitigate or overcome some common challenges in design and implementation, and potentially improve the reliability and usability of survey data.

Remote sensing tools can address the lack of census and other population data for the construction of sampling frames. Recent applications include the use of satellite imagery to construct a sampling frame in Mogadishu, Somalia, by stitching together satellite images and using image recognition algorithms to identify dwelling types and estimate the total inhabitable space and population. Potential biases remain, however,<sup>50</sup> and such data need to be rigorously ground-truthed.

### Box 12. High Frequency Surveys in South Sudan

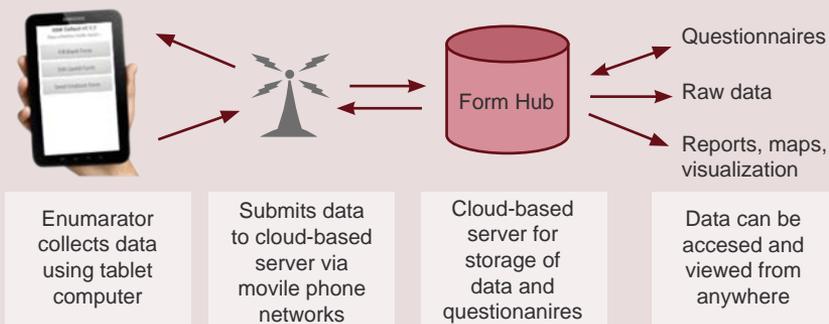
The World Bank and the UK Department for International Development (DFID) supported a High Frequency South Sudan Survey (HFSSS) to collect data on social and economic indicators from 2012–2013. The HFSSS pilot generated routine information across 4 of 10 state capitals<sup>51</sup> and was designed to reinforce the capacities of the new country's National Bureau of Statistics (NBS). It is expected that the HFSSS will be scaled-up to all 10 state capitals from 2014–2017 using a two-stage randomized design.<sup>52</sup> A key element of the tool is that data are “frequently” collected through household panels, weekly market prices, and daily exchange rates—thus improving the World Bank's capability to track changes over time and anticipate early warning signals.

- 
- 50 Potential biases cited by the authors include misclassification of image pixels, and—more problematically in an FCV context—the inclusion of empty buildings in population estimates, which could lead to over-estimates in areas heavily affected by conflict. See Jesse Driscoll and Nicholai Lidow, “Representative Surveys in Insecure Environments: A Case Study of Mogadishu, Somalia,” *Journal of Survey Statistics and Methodology*, Vol. 2, No. 1, 2014.
  - 51 These cities were selected by the NBS for the pilot based on their economic importance and population size, and included Juba in Central Equatoria, Rumbek in Lakes, Malakal in Upper Nile, and Wau in Western Bahr-el-Ghazal. Five rounds of the household survey were conducted in June and October of 2012 and January, June, and November of 2013, and covered both the rainy and dry seasons in South Sudan.
  - 52 The sample design for the household survey was a two-stage random cluster sample with the first stage being Enumeration Areas (EAs) and the second stage being households. The HFSSS used the EA sampling frame developed for the Fifth Population and Housing Census of Sudan carried out in 2008 as part of the Comprehensive Peace Agreement (CPA). Five EA's were randomly selected in the four state capitals for a total of 20 EAs during the first sampling stage. All households within these selected EAs were subsequently listed and eighteen households were randomly selected within each of these 20 EAs during the second sampling stage. This sample design produced a randomized sample of 90 households for each city or a total of 360 households for the HFSSS pilot.



(cont. box 8)

The HFSSS essentially entails the digitization of data collection and analysis, or what is commonly known as Computer Assisted Personal Interviewing (CAPI). The NBS employs four trained teams including one supervisor and two enumerators in each state capital who collect data on a daily basis.<sup>53</sup> The survey involves enumerators visiting households and markets using handheld tablet computers that run Android platforms and the Open Data Kit open software. The novelty of the CAPI survey is that completed survey forms are uploaded directly from the tablet via a mobile phone network to a free and open source cloud-based survey called Form Hub. This approach eliminates print paper survey forms, reduces enumerator error, eliminates data entry requirements, and makes data available in real time for visualization.



The HFSSS survey instrument features two components—a household survey and a market survey. The household survey questionnaire covers multiple sectors with modules on demographics, livelihoods and assets, health, education, food security and hunger, access to justice, government performance and service delivery, and safety and security. The market survey questionnaire collects weekly price and country of origin data for 20 retail and 8 wholesale products sold in local markets to track inflation. These products were selected by the NBS and constituted a subset of products from the South Sudan Consumer Price Index (CPI) that had high weights and were considered important for household wellbeing. These were primarily food items, petrol, and charcoal. The market survey also collects daily buying and selling rates of the South Sudanese Pound (SSP) to track fluctuations in the currency against the U.S. dollar.

53 The World Bank organized an initial training for the enumerators and supervisors on the survey methodology, questionnaire, and survey equipment in May 2012, with follow up refresher trainings held in August and December 2012 and May and September 2013 in preparation for subsequent rounds of the household survey. Daily exchange rate data and weekly market prices are being collected from May 2012 through December 2013.

The rapid spread of mobile phone technologies in even in the most fragile and unstable contexts has opened up new possibilities for data collection. This has taken two broad forms. While surveys—especially in low-income contexts—have traditionally been filled out using pen and paper, smartphones are increasingly used for data collection.<sup>54</sup> Mobile phone-based enumeration has a number of advantages: data encoding errors are often reduced; surveys can be automatically GPS-tagged, allowing for easy geolocation; data can be streamed and analyzed in near real time. There are also risks: tablets and cellphones can be a magnet for theft, particularly in areas with weak security. In some cases, the use of mobile phones, tablets, or GPS devices can raise concerns that enumerators are spying or violating privacy, which can raise safety concerns and potentially contaminate data collection. The benefits and risks of these tools should be carefully assessed before the survey is fielded.

While phone-based polls have been conducted for many years in high-income countries, they are starting to be employed in FCV contexts as well. For example, phone interviews have been employed to tap public opinion on constitution-making in Somalia, and text-message (SMS)-based surveys have been used to assess food security in Democratic Republic of Congo. Remote surveys have notable advantages: they can be deployed in insecure areas; they are significantly less expensive than in-person interviews; and they can be used for rapid polling in case of shocks or sudden change. That said, sampling considerations and potential for bias should be taken into account. Unlike high-income countries, there are rarely complete phone directories in FCV contexts, and not all populations of interest will have access to mobile phones. Text messaging, while effective for rapid surveys, may be best suited for short, simple questionnaires.<sup>55</sup> Increasingly, surveys may mix techniques, using traditional in-person interviews for longer, more complex questionnaires, and conducting rapid follow-ups remotely.

## Survey Ethics

Surveys pose ethical considerations that are different from other forms of data collection. Particularly in FCV contexts, surveys can potentially put participants and interviewers at risk. While it is easy enough to advise that surveys should do no harm, understanding the potential risks and tradeoffs is often more complicated in practice.

There are several key guidelines:

- **The safety of respondents and enumerators comes first.** It is important

---

54 A variety of software packages for mobile surveys exists, including many using open data platforms.

55 See Lucas et al. (2013).



to conduct continuous and active security assessments to identify whether interviewers can travel and administer the survey safely. Equally, it is important to assess whether asking or answering specific questions—or even just participating in a survey—might put respondents at risk, and whether steps need to be put in place to keep their identity anonymous.

- **Get informed consent.** Survey respondents have the right to know what they are participating in. This means providing basic information on what topics the survey will cover, how the data will be used and protected, and whether they will receive any benefits or face any physical or psychological risks. That said, there could be exceptions, specific conditions under which the broad social benefits of deceiving research participants outweigh the obligation to inform them fully.<sup>56</sup>
- **Participants' privacy and confidentiality should be protected:** their identity should not be released, and the information that they provide should not be disclosed. This may be particularly challenging in longitudinal surveys, which cannot be conducted anonymously, and in which respondents' personal information must be retained for longer periods.

It is important to note that there can be tensions between ethical obligations. For instance, in order for enumerators to work safely within a village, they may need to announce the survey to local elites. Once the interviewers' presence is known, it can be harder to maintain the anonymity and privacy of an interview. These are ultimately judgment calls that require careful balancing of benefits and risks. Ethical review processes can help identify and clarify these tradeoffs

## Partnership Arrangements

Although the World Bank sometimes directly implements surveys, **often Bank teams will partner with other organizations to collect data.** In many FCV environments, there will be a range of potential partners: international survey firms, local companies, and both local and international researchers. The capacity and relative benefits of each varies a great deal by context and by organization. It is important to recognize assumptions, and choose partners based on expertise, experience, and implementation capacity. Very generally, in low-income FCV contexts international actors will bring broader experience and technical expertise to the table than local survey organizations, but much less familiarity with local context. For this reason, if working with international firms or researchers, it can be helpful to facilitate partnerships between local and international survey groups.<sup>57</sup>

---

56 This assessment is not easy and should be made in concert with World Bank internal ethical review processes. If deception—either about the purposes of the survey or in a particular question—is used, it is critical to debrief a respondent at the end of the survey, to fully explain the nature and reason.

57 Partnerships can also provide a means for local capacity building. It is worth noting that many

If a survey is being put to public tender, the bid solicitation or terms of reference for a survey should include several key parameters, which can help ensure comparable and accurate bids:

- **Scale:** the number of respondents (or households), the geographic distribution (area, number of provinces, districts, or subdistricts)
- Proposed **sampling design**, including information on stratification or clustering
- **Pilot** requirements: number of pilots, number of respondents, location
- Arrangements for **translation** (and back-translation)
- **Interviewer selection**, including any requirements regarding team composition, experience, and training
- **Quality control** measures, including field supervision, and quality control for data entry and encoding if paper questionnaires are used
- **Deliverables**, including data format, interim and final reports, tabulations, and analysis, as required.

Finally, private survey firms are not the only partners. Governments will often be partners in supporting surveys, or key end users of survey data. **Some types of data can raise sensitivities with government:** perception and experiential questions on violence, corruption, and human and civil rights are notable examples. Even if a survey is independent, it is important to clarify expectations regarding what data will be collected, how the data will be used, and who will have access. Equally, it is important to work with national statistical offices and other government agencies that field surveys, to avoid duplication and ensure that the right issues are being covered.

## Where to Get Micro-Data

The number of outlets for micro-data is increasing and includes population surveys and instruments, as well as portals to engage with survey experts and practitioners. In collaboration with the International Household Survey Network, for example, the World Bank supports a micro-data databank of more than 1,300 household surveys in middle- and low-income settings.<sup>58</sup> The Empirical Stud-

---

international firms will work with local organizations as a matter of course, rather than building out local staff and infrastructure; when receiving tenders for survey services, it is critical to get details on who firms are contracting with, their level of capacity, and the relative margin.

58 At present there are comparatively few dedicated surveys covering just FCV countries. For example, the World Bank library includes just 162 surveys. Of these, there are two surveys each for Guinea and West Bank and Gaza and one survey each for Iraq, Mali, and Nepal. The database does not have surveys for the other 28 countries and 1 territory on the 2013 fragile situations list.



ies of Conflict project has a data portal that pulls together public opinion surveys, census information, and administrative datasets for a subset of fragile and conflict-affected states. The Complex Emergency Database (CE-DAT) has aggregated over 3,000 health, service, and nutrition surveys, including many from FCV areas.

Meanwhile, organizations are also working to generate new, internationally comparable data on FCV contexts. Notable among these efforts are the Strategy for the Harmonization of Statistics in Africa (SHaSA) survey pilots; the Secure Livelihoods Research Consortium; efforts by the World Bank to generate comparable data on public sector corruption; and over 250 health and demographic surveys produced across 90+ countries by Measure DHS.

Good examples of groups seeking to promote the generation and dissemination micro-level household survey data in FCV include the Micro Level Analysis of Violent Conflict (MICROCON) project and Households in Conflict Network (HiCN). The latter features useful partners, as well as research generated from micro-data. Lastly, new online discussion boards are linking practitioners and providing resources on designing and implementing surveys in developing country contexts.

## Additional Resources

Bradburn, N., Rips, L., and Shevell, S. 1987. "Answering Autobiographical Questions: The Impact of Memory and Inference on Surveys" *Science* vol. 236.

Brück, T., Justino, P., Verwimp, P., and Tedesco, A. 2013 "Measuring Conflict Exposure in Micro-Level Surveys" Washington, DC: World Bank.

International Household Survey Network (ISHN). 2009. See especially the Survey Quality Assessment Framework.

Krosnick, J., and Presser, S. 2010. "Question and Questionnaire Design" in Marsden, P. and Wright, D., eds., *Handbook of Survey Research*, Emerald Group Publishing.

Johnson, N., Spagat, M., Gourley, S., Onnela, J., and Reinert, G. 2008. "Bias in Epidemiological Studies of Conflict Mortality" *Journal of Peace Research*, Vol. 45, No. 5.

Lavrakas, P. 2008. *Encyclopedia of Survey Research Methods*, Sage.

Lucas, H., Batchel, S., and Berdou, E. 2013. "Real Time Monitoring and the New Information Technologies," *IDS Bulletin*, Vol. 44, No. 2.

- Parks, T., Colletta, N., and Oppenheim, B. 2013. *The Contested Corners of Asia: Subnational Conflict and International Development Assistance*. The Asia Foundation, pp. 29–30, 155 (fn. 70).
- Petra, L. 2008. "Questionnaire Design in Attitude and Opinion Research: Current State of an Art" Jacobs University, Working Paper 655, Bremen.
- Schulhofer-Wohl, J. 2014. "Armed Group Repertoires and Recollection in Survey Research" Households in Conflict Network Working Paper 171, March.
- Survey Research Center. 2010. "Guidelines for Best Practice in Cross-Cultural Surveys." Ann Arbor, MI: Survey Research Center, Institute for Social Research, University of Michigan.
- United Nations Department of Economic and Social Affairs. 2005. Statistics Division "Household Sample Surveys in Developing and Transition Countries." New York: UN.
- Vera Institute of Justice. 2003. "Measuring Progress toward Safety and Justice: A Global Guide to the Design of Performance Indicators across the Justice Sector" New York: Vera.