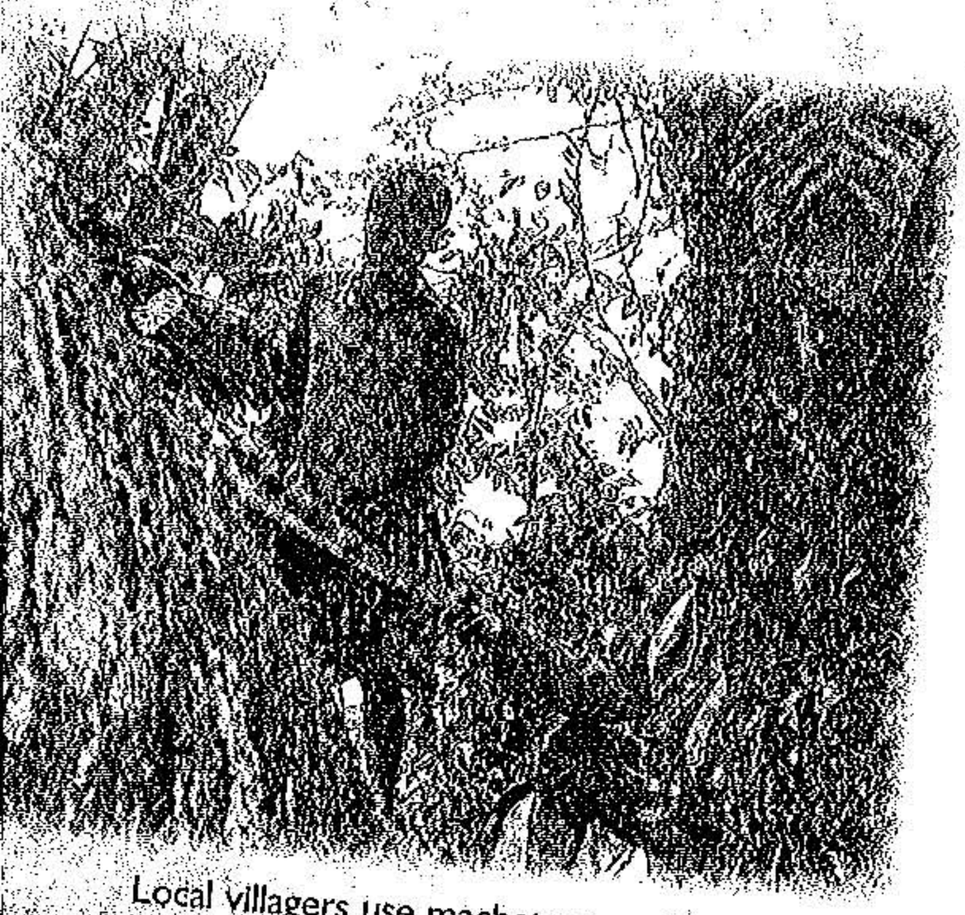
Collecting **Data** in Challenging Settings

In the Global South—say, East Timor—data collection is not for the faint of heart or weak of stomach.

Jana Asher

Truck wreck in Sierra Leone. Good thing the truck was big, or the pothole might have swallowed it whole.



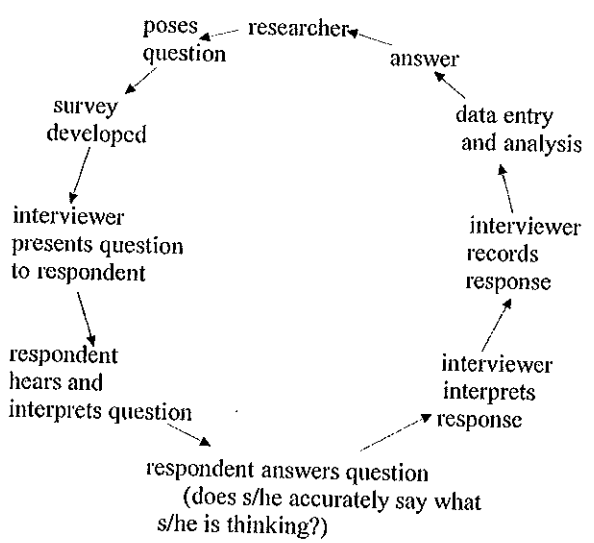
Local villagers use machetes to remove a tree trunk from the road in East Timor. There were no governmental services available to clear the tree; the villagers were responsible for keeping their own passages off the mountainside clear.

It is very early in the morning when water falling on my face jolts me awake. The raindrops are blowing into the small bedroom through a gap between the roof and side wall, and, unfortunately, the platform bed on which I am curled is occupied by four other women. Having nowhere to move, I coil up into a tighter ball under our shared blanket and attempt to sleep . . . despite the small pool of liquid that has formed a sheen on my cheek.

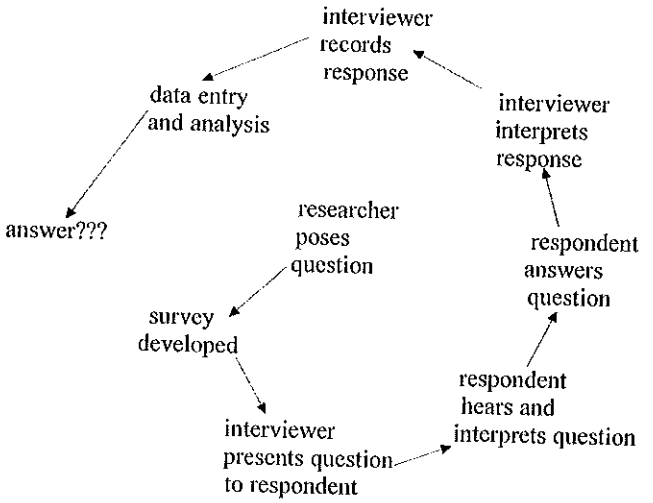
The next morning, the occupants of the small village in which we are stranded will use machetes to cut out a portion of tree trunk that is blocking our path down the mountainside. We will make it about one-third of the way down the mountain in our truck before we are blocked by a second tree trunk, bigger than the first, and are therefore forced to abandon our vehicle. We will walk the rest of the way down the mountain in a state of hyper-alertness, listening to the creaking of the trees and hoping that one of them does not decide to fall on top of us. In the evening, we will finally make it to Dili, the capital of East Timor. That night, I will call my irate and somewhat despondent husband and assure him I am fine, despite having missed our planned phone call the night before. And then I will collapse in my warm hotel room bed, half a world away from my family and friends.

To be fair, when we started out that morning, the weather was hot and dry. By an accident of fate we just happened to select a village on top of a mountain as a test site for our questionnaire on the very day that the rainy season decided to arrive. We completed our testing in a thunderstorm, and by the time we were done, the path we had taken up the mountain no longer existed, having been subsumed by a tree with a diameter greater than a man's height.

All for the sake of testing a questionnaire.



Well-designed Data Collection



Poorly-designed Data Collection

Figure 1. The data collection communication process

Data collection in the Global South is not for the faint of heart or weak of stomach. Nor is it for the overly rule-bound or uninventive. It requires an intimate understanding of your own fragility and mortality, a sense of adventure, a respect for the knowledge and wisdom that each person you encounter provides, and a limitless supply of patience. Finally, it requires a desire to bring the highest intellectual and scientific rigor to the most difficult of circumstances ...coupled with the understanding that you will never truly succeed at doing so.

But we are getting ahead of ourselves. Let us start by discussing why the questionnaire design and data-collection processes are so essential to the quality of the resulting data, and then review the ingredients that constitute good questionnaire design and data collection. Then we can return to East Timor—by way of the Middle East and Africa.

What Can Go Wrong?

It seems like a deceptively simple process, one perhaps that you first tried in an elementary or junior high school class. You have a research question that can be answered by survey data, so you write some questions, slap them onto a form, copy, distribute, collect, and presto: instant data!

Well, yes and no. Any data-collection process represents a complex chain of communications, and as with any other chain, one weak link can break it. The more complicated the data-collection process, the more links in the chain, and the more opportunities for communication to break down. For a traditional interviewer-administered survey, multiple individuals must have a nearly identical conceptual understanding of a question, as shown in Figure 1.

When everything goes well, the chain is a circle as presented on the left: That is, the interpretation of what information is desired is identical across the many individuals in the chain, leading to accurate reporting and recording of information. What can happen, however, is that the chain forms an outward spiral like the one on the right—and the information collected is not identical to that which was originally desired by the researcher.

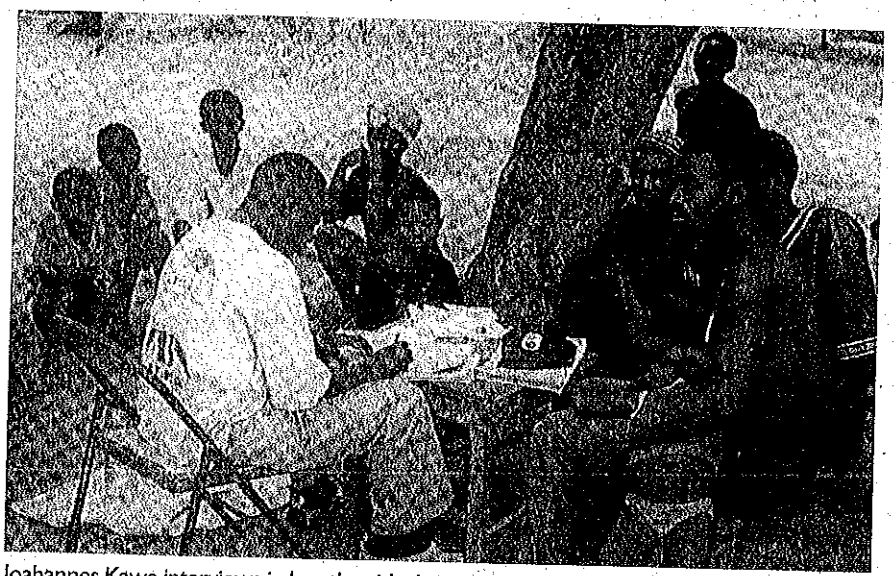
Two main issues that can arise in the process of communication are misinterpretations (errors in the communication between people) and inaccuracies in transmission (such as error caused by the interviewer writing down the wrong code on the survey form). The misinterpretation type of error is best explained by examples. Let us start with a basic question that is fraught with interpretation issues and would therefore almost never appear on a pretested questionnaire.

Avoiding Ambiguous Questions

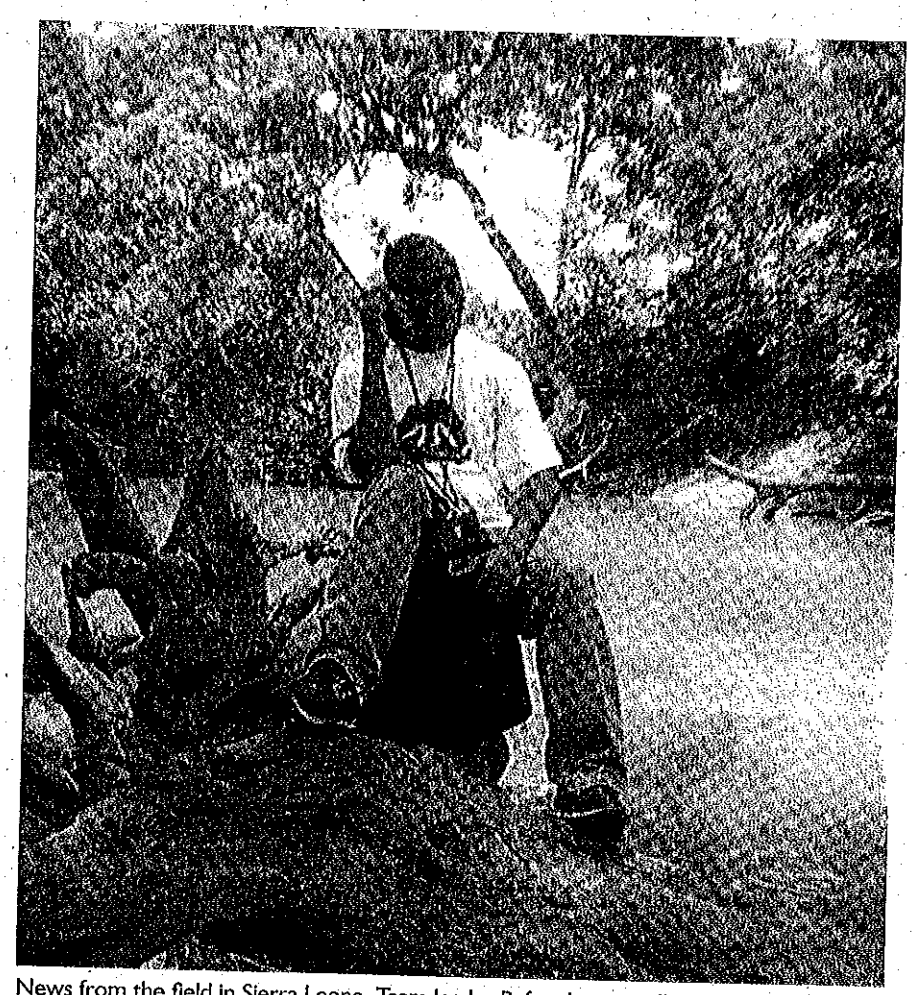
To a critical observer, there are obvious ambiguities in this question. Within what time period do we mean? Do we mean household or individual income? Reported taxable income? Are alimony payments, gambling winnings, and other types of income included? What about bartered items? How is a respondent to know?

Here is another real-life example: the testing of a web-based survey that was to be administered to expatriate Iraqi physicians so that they could describe their professional and other experiences in their new countries of residence. The research group that hired us to test it had been advised by a well-trained and highly reputable Iraqi physician that all Iraqi medical schools were conducted in English. Therefore, the survey could be administered in English and the doctors would have little difficulty understanding the questions.

Unfortunately, he was wrong. As an illustration, note the following



Joahannes Kawa interviews in less than ideal circumstances. Note that privacy is not a common commodity in the rural areas of Sierra Leone, where the arrival of the interviewing team was considered a major event.



News from the field in Sierra Leone. Team leader Bafara Jawara calls via satellite phone to report on progress in the Bonthe district.

transcription from a pretest interview. The volunteer was instructed to read the question aloud and then tell the interviewer all thoughts he had as he determined his answer to the question. The goal of the question was to determine whether the physician was required to become recertified to practice medicine in his new country.

Volunteer interviewee (reading): If you had to go through a credentialing process or are currently in a credentialing process in your new country, how many years did it take? Please do not answer if you did not need to go through a credentialing process in your new country.

Volunteer interviewee (responding): Do you know if the credentialing process is the asylum process or the red card or green card?

Interviewer: I'm not allowed to say anything, because it has to be how you interpret it.

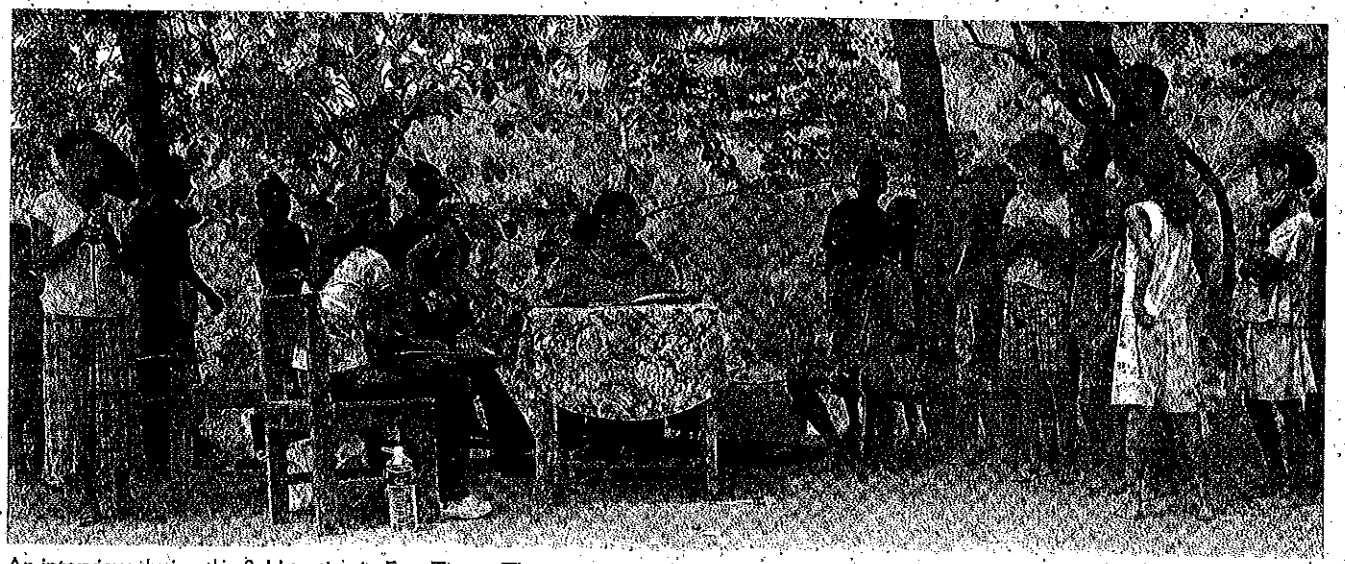
Volunteer interviewee: Yeah, I need to understand what means ...This is asking me...like through the embassy or through the ...I don't understand what mean [sic] "credentialing." (Looks it up in English-Arabic dictionary.) Here credential is like ambassador or delegation ...yes. What mean this [sic]?

Interviewer: So how would you answer this question?

Volunteer interviewee: This is paper of ambassador of delegation [sic] ...I don't know.

You do not need to be a seasoned questionnaire designer to see a big problem in the respondent's interpretation. We and our clients determined that their "cultural expert" on Iraqi physicians was familiar only with the "elite"—that is, those physicians who attended the best Iraqi medical schools and held the most prestigious positions in their home country. In fact, of the four Iraqi physicians we interviewed during the pretest of the questionnaire, only one came close to interpreting all the questions accurately.

As a result of the pretest, the research team members developed an Arabic version of the questionnaire, thus saving themselves from collecting data that would be, for all intents and purposes, garbage. In general, a rigorous testing procedure will be required to ensure that a survey collects the data it intends



An interview during the field testing in East Timor. The interviewer, Jacinta Gonsalves, is sitting at the table; her supervisor, Silvia Verdial da Silva Lopes, is sitting to her right; and the respondent is the man sitting to her left. Note that privacy was virtually impossible during the interviewing process.

to collect—in other words, that the data collection process ends as a circle instead of a spiral.

Best Practices for Data Collection

So what is the right way to design and test a questionnaire and interview respondents? Well, the short answer is that it depends, but there are some best practices that have been developed over the past few decades. Here is a look at the process of developing the questionnaire and training interviewers to administer it.

Testing the questionnaire

Testing a questionnaire requires multiple steps. A first round of testing should include review of the survey instrument by at least three different types of experts: an expert on the subject of the survey, an expert on the population to be surveyed, and an expert on questionnaire design. Although those roles might overlap—for example, your questionnaire design reviewer might be the same as your expert on the population to be surveyed—none of those experts should be the person who initially designed the questionnaire. Even the most experienced questionnaire designers make mistakes.

Once that initial review is completed and the questionnaire has been modified based on the expert comments, the real testing begins. There are several possibilities for the next round of review. One of the most successful techniques is called "cognitive interviewing." Cognitive interviewing allows the survey designer some insight into the thinking processes of the respondent.

During a cognitive interview, two trained professionals administer the survey to a test respondent. One of the professionals serves as the interviewer, and one records his/her impressions of the interaction between the interviewer and respondent, including tone of voice and body language that might indicate confusion or a strong emotional reaction. The respondent is asked to "think aloud"—that is, to verbalize his/her thoughts while responding to the question. In addition, the interviewer might ask probing questions—questions about the respondent's interpretation of or thinking about the survey.

Those probing questions might be either ad-lib or carefully developed prior to the cognitive interviewing process. If possible, the cognitive interview will be tape- or video-recorded for study later. As an example, the transcription of the interview of the Iraqi physician given earlier in this article was taken from a cognitive interview.

Cognitive interviewing can occur as a single process or in waves. When cognitive interviewing occurs in waves, between each wave the survey developer modifies the survey on the basis of what was learned during the previous wave. Following the example of the question "What is your income?" we can imagine a first round of cognitive interviews might uncover the issue that there is no time frame given. The survey designer might respond by reformulating the question as "What was your income over the past 12 months?"

A second round of cognitive interviews might reveal that some respondents are including interest income and some are not. The question might then be reformulated as "What was your income from gainful employment over the past 12 months?" and so on, until the cognitive interviews indicate that the respondents to the question are interpreting it as the survey developer intended.

Limitations of cognitive interviewing

One issue with the cognitive interviewing process is that it is time and labor intensive, so only a limited number of cognitive interviews can be completed during any wave. Therefore, it is a good idea to perform a final field test after the cognitive interviewing. A field test is a small run of the fieldwork for the survey, after which the results are tabulated to make sure that they seem appropriate. The field test allows both input from a large number of individuals from the population of interest and also an opportunity for the interviewers to practice with the survey before the real fieldwork begins. One hopes that any problems found at this stage will be minimal and

A team in Sierra Leone reviewing questionnaires under the supervision of team leader Mohamed Daboh. From left to right: Andrew Simbo, Mohamed Daboh, Antoinette Licon, Nancy Joseph, and Joahannes Kawa. Note that the team is working by candlelight; electricity in the field was a luxury, as were flush toilets. Also note the spiffy Joint Statistical Meetings bags, donated by the American Statistical Association.

Transportation in the Bonthe district of Sierra Leone. The team members that covered the Bonthe district spent the majority of their mission being transported by boat. For this reason, they were issued life preservers and special containers in which to store the survey forms and their equipment and personal belongings. Here team members rest as they travel to the next village. The Bonthe team interviewed in the remotest village sampled for the survey. It was on an island that belonged to the district of Pujehun, and the Pujehun team had been unable to reach it. The Bonthe team members, led by Bafara Jawara, were required to ride in a boat for 16 hours and then hike for 10 miles to get to the village.

easily corrected, and the survey will be ready to commence in earnest.

Another issue that frequently arises during questionnaire design is the need to translate the questionnaire into one or more languages. The base minimum for translation is a combination of forward and backward translation—that is, one individual performs a translation between the original and target language, and another individual translates it back into the original language. The pretranslation is then compared to the version that has been forward- and back-translated to find inconsistencies, and the translation is then corrected.

Although this has been an industry standard, recent research has suggested that a more rigorous technique must be used, especially in the case of a questionnaire that is being translated into multiple languages simultaneously. One option is to develop the questionnaire in the multiple languages at the same time, with individual teams for each language starting from a base set of concepts. In that case, cognitive interviewing and other testing methods will occur in all languages, not just the base language.

Training the interviewers

An essential aspect of the questionnaire design and testing process is the appropriate training of the interviewers. In most large survey projects in the United States, the interviewers are trained to "stick to the script"—in other words, to read the questions on the survey instrument exactly as written, to follow the questions in the order given, and not to offer additional explanation of the questions unless it has been preapproved by the survey manager.

This protocol is designed to minimize interviewer bias in the answers—that is, changes in how respondents will answer questions that are due to some behavior of the interviewer. Interviewers may inadvertently change the meaning of a question if they do not read that question exactly, or they may cause respondents to favor one response over another due to the interviewer's clear preference for that response. For that reason, interviewers are also taught to present the questions in a neutral way and to not implicitly or explicitly express their own opinions as to an appropriate answer.

Interviewers must learn the appropriate way to administer the survey. However, there are other aspects of interviewer behavior that must be addressed, including appropriate voice control and body language. Interviewers are trained on techniques for building rapport, asking sensitive questions, and maintaining the confidentiality of the responses of the individuals interviewed. Finally, interviewers need to understand methods for keeping themselves safe in the field.

What Makes Survey Practice Different in a Developing or Transitional Country?

Much of our understanding of what constitutes good survey practice has grown out of research that has taken place in the United States, Canada, and Europe—all vastly different environments from those of the Global South. There are several reasons why the data-collection methods discussed above might need significant alteration to be useful in the developing context.

A particularly pernicious problem—one that just recently has become a research priority of government statistical organizations in the developed world—is the need to administer a survey to a diverse population comprised of multiple ethnic groups that speak a diverse set of languages. In the past, questionnaires often have been developed without consideration of how particular questions will be interpreted or understood across cultural groups, or whether particular concepts are directly translatable at all: an issue that the forward-/back-translation industry standard does not adequately address.

Populations in the Global South might be more varied than those in the Global North for other reasons—including varying levels of literacy and different understanding and tracking of time (e.g., reliance on agricultural cycles rather than a Gregorian calendar). In the context of sensitive questions, ensuring privacy might be close to impossible in the Global South context, where entire families might share a one-room home or apartment and the arrival of interviewers is a villagewide event. And the cultural preferences of many Global South countries lead to higher rates of "acquiescence bias": the higher likelihood of respondents answering yes to a question, not because the affirmative is true but because they want to please the interviewer, who is perceived as being in a position of authority.

Are There Best Practices for Data Collection in the Global South?

The answer to this question is yes and no. There are several organizations—including the World Bank and United Nations—that have compiled the state of the art in random sample surveys in the Global South. However, there is still much research to be done as well as ongoing issues with the indiscriminate transfer of data-collection methodology.

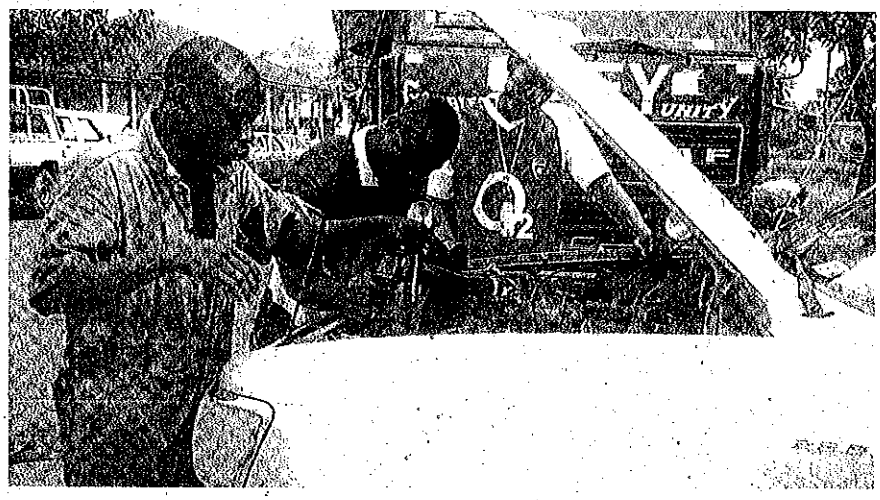
Cognitive interviewing in the field in East Timor. From left to right: Jana Asher, Duarte da Silva, and the survey respondent with her daughter.

One important difference between fieldwork in the Global North and Global South is the availability of support infrastructure in the field. Common developed-world conveniences like cellular phone networks or landlines, hospitals and clinics, restaurants and hotels, and regular electricity and running water simply might not be available in some parts of the Global South. Teams of interviewers can be outfitted with satellite phones and carry their own medical supplies into the field. Interviewers can be vaccinated against local diseases and carry food and shelter to remote locations. And plenty of matches and candles will allow interviewers to complete their activities after dark.

In addition, interviewers must be prepared in advance for difficulties such as flat tires, large potholes, or the need to forget the car or truck altogether and travel by motorcycle, boat, or foot. And you never know when your vehicle will not be able to proceed because there is a fallen tree in your way!

What about the design of the questionnaire in the Global South context? In this researcher's experience and opinion, in Global South countries where multiple tribes or cultures may be part of the sampled population, training the interviewer to recite the questionnaire verbatim, with no ability to explain if the respondent is confused, does not lead to quality data. In fact, a reluctance to adapt to the interviewee can harm the interview process. Rather, the interviewer needs greater training and greater ability to improvise in the field. Additionally, although they are not often used, techniques like cognitive interviewing and more recently developed language translation techniques can and should be used in the Global South.

Many cultures in the Global South are more attuned to agricultural cycles and timing than Gregorian divisions of time. In many cultures, even significant dates, like one's birth date, are not known or not important. That can significantly impact the quality of surveys that require recall of events, unless specific techniques are developed for those populations. A technique called the "calendar method" has shown great promise. The interviewer assists the respondent in developing a written calendar of important "landmark" events in his/her life to aid recall of events asked about during the interview. However, in its typical form, the calendar method requires the respondent to be literate and familiar with the Gregorian calendar. The following section describes an alternate method of assisting illiterate respondents in recalling date information: a



Team leader Mohamed Daboh watches as a local mechanic in Bo Town attempts to fix his team's car during fieldwork. Car problems plague fieldwork in a country like Sierra Leone, where there is insufficient infrastructure to maintain the roads—or even pave the majority of them.

Sahr Gbondent holds up the remains of an interview team's tire during fieldwork in Sierra Leone. During that project, each vehicle was equipped with two extra tires for each two-week mission. Very often, they weren't enough.

method that the author developed in her fieldwork.

A Small Research Example

The survey was a national survey of human rights violations that occurred during the armed internal conflict of Sierra Leone. The U.S. State Department sponsored the survey as part of a general program of documenting war crimes. The participating organizations were the American Bar Association and Benetech. All three groups wanted the data to be used by the Sierra Leone Special Court during the prosecution of war criminals.

For that to happen, each violation needed to be associated with a perpetrator group and a date. In addition, we were asked to collect age of the victim at the time of violation, the duration of the violation (including duration until death if appropriate), and the current age of all members of the respondent's household. However, a large percentage of the population of Sierra Leone is illiterate and does not regularly consult the Gregorian calendar of years, months, and weeks. Rather, the people are attuned to the seasons (rainy and dry) and other important events such as religious holidays and school terms.

Our solution was twofold: We provided the interviewers with more latitude to probe the respondent for information (not sticking to the script of the questionnaire), and we crafted a series of prescripted probe questions that asked the respondents to determine whether a violation occurred before or after a national date of importance.

Although the first decision deviated from standard survey practice regarding interviewers, we felt that the complexity of the information sought required that interviewers have the ability to be more creative in the field, and that the additional information elicited far outweighed the potential interviewer bias introduced.

The second decision, however, rested firmly on current understanding of memory and cognition by psychologists. Current theory suggests that our recall of date information is based on a few "landmark" events for which we have memorized the date on which the event occurred, combined with the storage in memory of events in series form. In other words, an interviewee might remember that her birthday is October 17 and then

Table 1—Time Probe Results for Sierra Leone War Crimes Documentation Survey

Probe	Victim Age During Violation	Violation Start Date	Violation Duration	Death Duration	Resident Age
No probe needed + 1 probe	766	171	85	20	701
No probe needed	44,168	11,040	57,488	10,018	32,013
Missing value	3,229	2,107	4,899	370	1,825
1 probe code recorded	16,528	32,538	2,217	380	4,310
2 probe codes recorded	26	18,577	28	0	1
3 probe codes recorded	1	283	0	0	0
4 probe codes recorded	0	1	0	0	0
More probe codes	0	0	0	0	0
Total probe events	**19,784**	**51,399**	**2,245**	**380**	**4,311**

remember that about two weeks later she had her furnace serviced—but it is very unlikely she will remember that her furnace was serviced on October 29.

In the case of the Sierra Leone survey, we created a probe question for each of eight events of national importance that were spread through time in a way that allowed interviewers to determine the year in which a particular respondent's reported human rights violation had occurred. For example, the probe "Did that happen before or after the invasion of Freetown?" allowed the interviewers to determine whether an event occurred (roughly) prior to 1999, during 1999, or later.

In addition, the interviewers were provided with several scripted probes to help them narrow the time frame during which the violation could have occurred. Those probes referenced seasons (rainy versus dry), religious events (before or after Christmas or Ramadan), school terms (before or after a particular term started or ended), and, if needed, age of the respondent (both age of the respondent when the event happened and age now). The interviewers were free to use any combination of probes they felt was warranted, but they were required to record which probes they used (or note if they had not used any). The strong advantage of this method versus more traditional calendar-based methods was its ability to be used with illiterate individuals.

Did it work? The evidence given in Table 1 would suggest so. Using the recorded information about which probes were used to ask about which violations, we determined that 79.4% of violation dates relied on probing for the date to be determined. Events that were easier to remember were durations—as given by the results for violation duration and time from violation until death—only 3.5% of those cases required probes.

Note that typically only one probe was required to determine dates. We believe that respondents, once introduced to the probing method, engaged in self-probing to recall ages and dates later in the interview. This theory is supported by observations of the cognitive interviewing for the survey, during which respondents began to self-probe. Please also note that the use of landmark-based probe events is highly supported by the response rate to time-related questions of the survey: Response rates were very high for resident age (99.9%), victim age (99.9%), violation start date (99.9%), and violation duration (96.9%).

Summary

Many people required to analyze data do not see those data until they are nice and neat, either in a spreadsheet or another organized format. They might not realize how much error can be introduced through the data-collection process if the process is not carefully planned. You now know what can and will go wrong when appropriate questionnaire design and data collection protocols are not followed—or even when they are.

Scientists are hard at work determining new and better techniques for eliciting information from a variety of people across the world. So the next time you are given a data set to analyze, be sure to ask how the data were acquired. You might be surprised by how little or how much was done to produce data of the highest quality possible—whether or not the questionnaire designer ended up stuck on the top of a mountain. ■

Further Reading

Bulmer, M., and Warwick, D. P., eds. 1993. *Social research on developing countries: Surveys and censuses in the Third World*. London: University College London Press.

Casley, D. J., and Lury, D. A. 1981. *Data collection in developing countries*. London: Oxford University Press.

Tourangeau, R., Rips, L. J., and Rasinski, K. 2000. *The psychology of survey response*. Cambridge, UK: Cambridge University Press.

United Nations. 2005. Household surveys in developing and transitional countries. *http://unstats.un.org/unsd/Hhsurveys/* (October 30, 2009).

Willis, G. B. 1999. Cognitive interviewing: A "how to" guide. *http://appliedresearch.cancer.gov/areas/cognitive/interview.pdf*.