Data Collection: General Considerations

Julie de Jong, 2016
(2010 Version: Beth-Ellen Pennell, Rachel Levenstein, and Hyun Jung Lee)

Introduction

Collecting comparable data in the context of multinational, multiregional, and multicultural surveys (“3MC” surveys), is a highly complex task, in which one can expect to encounter a variety of languages and cultural contexts. Even in a single country, the target population may not be linguistically, ethnically or culturally homogenous. Such cultural heterogeneity could manifest itself through a wide variety of dimensions that could impact data collection efforts. For example, languages spoken may not have a standard written form, or respondent literacy rates may be vastly different. The geographic topography may be difficult (e.g., remote islands, deserts, or mountainous regions), and weather and seasonal impediments (e.g., winter/summer, monsoons), national and religious holidays (e.g., the Christmas season, Ramadan), or political upheavals may make the harmonization of fielding times across different countries impractical. Moreover, some populations may be inaccessible because of migration patterns or interviewer safety concerns, or they may be only accessible under special circumstances (e.g., miners in camps, or populations in which part of the population goes on long hunting or fishing trips).

Countries also vary widely in both their survey research infrastructures and in their laws, norms, values, and customs pertaining to data collection and data access. Certain modes of administration may be inappropriate or not feasible in some situations. In addition, the size and composition of nonresponse will likely vary due to differences in cooperation and ability to contact respondents. Some countries officially prohibit survey research (e.g., North Korea) or severely restrict data collection on some topics.

While a survey conducted in a single country might face one or more of the challenges mentioned above, the probability of encountering multiple hurdles is much higher in a large-scale 3MC study. What is atypical in the one-country context often becomes the norm in 3MC contexts. Moreover, the assumed homogeneity and common ground that may, broadly speaking, hold for a single-country study contrasts with the obvious heterogeneity of populations, languages, and contexts encountered in multinational studies. Because of the heterogeneity of target populations in cross-cultural surveys, allowing some flexibility in data collection protocols can reduce costs and error.

In some cases, a coordinating center dictates data collection decisions across all countries involved. The European Social Survey (ESS), for example, mandates the mode in each country, while the International Social Survey Programme...
(ISSP) allows a certain amount of flexibility. See Study Design and Organizational Structure for more details.

These guidelines are intended to advise data collection decision-makers as they consider the issues and requirements relevant to different data collection modes and provide extensive recommendations for the practical implementation of data collection in different modes. Because the guidelines and lessons learned vary greatly depending on the specific mode of data collection, we begin with general considerations relevant for data collection in any mode and then provide further guidelines and lessons learned in three subsequent chapters for the main modes of data collection used for 3MC surveys as follows:

Data Collection: General Considerations (these guidelines)
Data Collection: Face-to-Face Surveys
Data Collection: Telephone Surveys
Data Collection: Self-Administered Surveys

For a discussion of the advantages and disadvantages of specific modes, key factors involved in mode choice, and whether to standardize mode across locations, see Study Design and Organizational Structure.

Because difficulties in data collection can be extreme in countries where infrastructure is limited, these guidelines heavily emphasize the challenges of data collection in such contexts.

Guidelines

Goal: To achieve an optimal cross-cultural data collection design by maximizing the amount of information obtained per monetary unit spent within the allotted time, while meeting the specified level of precision and producing comparable results.

1. Before beginning fieldwork, assess the feasibility of conducting the research in each target country and culture.

Rationale

Local knowledge can be critical to understanding cultural traditions and customs, possible limitations, and the feasibility of the research. Experienced researchers, interviewers, and key stakeholders familiar with the topic or the population under study can help assess concerns and challenges and suggest potential solutions.
Procedural steps

1.1 Assess the appropriateness of (1) the constructs to be studied, and (2) the mode of data collection selected (van de Vijver & Leung, 1997). For detailed information about different data collection modes, see Data Collection: Face-to-Face Surveys, Data Collection: Telephone Surveys, and Data Collection: Self-Administered Surveys.

1.2 Gather information from the coordinating center on major survey design features. These might include the survey topic and questionnaire items, intended mode of administration, instrument technical design, respondent burden (e.g., length of interview, complexity of topic), and proposed methods for dealing with nonresponse.

1.3 Gather information from people who are familiar with data collection in the area and from people who may not be familiar with survey data collection but who are familiar with, represent, or may share characteristics with the population of interest. If possible, conduct focus groups and one-on-one interviews with individuals within the contracted survey organization and others who have previously collected data within the country or location.

1.3.1 Solicit the help of local collaborators or researchers. Local collaborators may have a solid understanding of relevant cultural concerns or challenges or they may be able to help gather information from other local individuals who are more familiar with data collection and the population(s) of interest.

- Provide local collaborators or researchers with a detailed description of the protocol, including the proposed mode of data collection, nonresponse reduction techniques, timing, interviewer training, remuneration, monitoring, and the general framework for data collection.
- Explain and clarify any survey terminologies to ensure common understanding.
- Request feedback on all aspects of the proposed study.
- Arrange to be present (even if by phone or other means of communication) when local collaborators are collecting information from local resources to clarify and probe when needed. However before making a decision to join those meetings, assess whether participating in those meetings might make locals uncomfortable and wary of providing information.

1.3.2 Elicit information from these local human resources and any relevant administrative bodies on:
Data Collection: General Considerations

1. Population issues (e.g., local knowledge about the survey, family structure and household listing issues, literacy levels, unwritten languages and cultural norms).

2. Logistical issues (e.g., seasonal accessibility, locked dwelling units, secured or dangerous areas, and connectivity issues).

3. Issues related to mode choice (see Study Design and Organizational Structure, Data Collection: Face-to-Face Surveys, Data Collection: Telephone Surveys, and Data Collection: Self-Administered Surveys).

4. Issues related to interviewers if an interviewer-administered mode is used (e.g., availability of interviewers, background, and safety concerns).

5. Human protection issues (e.g., legal and cultural permissions which may be necessary to conduct the study) (see Ethical Considerations).

**Lessons learned**

1.1 While outside input is often helpful, recognize that negative feedback may, in part, reflect uncertainty rather than concrete obstacles. Such feedback can, however, alert researchers to constraints that require attention. For example, in an early survey of mass media communication behavior in the Middle East, experts predicted that data collection would not be possible in Arab countries because the experts believed the populace would think that the interviewers were agents of the government. The experts also suggested that women could not be hired as interviewers and that it would be impossible to survey previously unsurveyed groups, such as the nomadic Bedouin tribes. The research team, however, was successful in their data collection efforts (Carlson, 1958).

1.2 While a mixed-mode design can reduce the cost of data collection by allowing for increased flexibility to accommodate local contexts, it may also create an additional layer of complexity and, thus, the overall costs for the subsequent harmonization of data by coordinating centers. The Gallup World Poll implements a mixed mode design in which the telephone is used in countries where 80% or more of the target population is covered and face-to-face interviewing is used in countries with lower telephone coverage. The reported costs of telephone surveys are much lower than face-to-face modes (Biemer & Lyberg, 2003), so overall data collection costs are reduced. However, comparability problems due to different modes (phone in one country, face-to-face in another) may exist (Gallup, 2007).
1.3 In a cross-national context, the impact of mode can be confounded with cultural differences. For example, when the International Social Survey Programme (ISSP) began, the required mode was a self-administration. However, low literacy levels in some countries necessitated the use of interviewers. Both response rates and reports from substantive measures differed widely, possibly as a result of differences in mode (Skjåk & Harkness, 2003). Therefore, reported variation between countries on survey estimates may indicate substantive differences or may be a result of mode effects and interviewer effects.

2. Decide whether the desired information can best be collected by combining qualitative methods with the standardized survey.

**Rationale**

A mixed method data collection approach can increase data quality and validity in a number of ways.

First, applying a combination of research methodologies to study the same phenomenon facilitates the validation of data through cross verification, while each method counterbalances the potential limitations of the others (Hulme, 2007). Qualitative and quantitative data collection and analysis methods can be used iteratively to strengthen both approaches. For example, qualitative, less structured interviews may permit a more positive interaction between the interviewer and the respondent increasing the accuracy of the information the respondent provides as well as his or her willingness to provide such information. Qualitative methods can also place the behavior of respondents into a broader context and can improve data coding by revealing unanticipated influences.

Second, mixing qualitative and quantitative methods can address the complexity of sensitive topics or cultural factors more fully than can quantitative methods alone (Bamberger, Rao, & Woolcock, 2010). Finally, it is not necessary to draw a strict dichotomy between qualitative and quantitative approaches; researchers may remain open to switching between the two so-called paradigms within the course of a study (van de Vijver & Chasiotis, 2010).

**Procedural steps**

Choose data collection methods to fit the aim of the research question (Axinn & Pearce, 2006).
2.1 Consider combining less structured interviewing, field notes, observation, historical materials, or event history calendars with the standardized survey (Axinn & Pearce, 2006).

2.1.1 In the social sciences, the term "methodological triangulation" is often used to indicate that more than two methods are used in a study to double (or triple) check results (for further information on methodological triangulation and integrating qualitative and quantitative methods in data collection, see Further Reading).

2.1.2 Triangulation can also widen and deepen one’s understanding of the phenomenon being studied.

2.2 Ethnosurveys offer an approach that combines survey and ethnographic data collection and allows each method to inform the other throughout the study. Equal weight or priority is given to each method. Quantitative data is collected in a manner that falls between a highly structured questionnaire and a guided ethnographic conversation, which is helpful in contexts where rigid structure may be inappropriate but where some standardization is needed for comparison purposes. See Massey (1987) on the theory and practice of ethnosurveys.

2.2.1 Determine whether your study is retrospective, prospective, or both. Calendar methods are more efficient for retrospective studies while longitudinal designs are more efficient for prospective studies (Axinn & Pearce, 2006; Freedman, Thornton, Camburn, Alwin, & Young-Demarco, 1988).

2.2.2 Remember that traditional qualitative methods can be more expensive and time consuming than a standardized survey (Massey, 1987; Morse & Niehaus, 2009).

Lessons learned

3MC projects have successfully combined qualitative and quantitative methods of data collection in many different ways.

2.1 The Tamang Family Research Project, conducted in Nepal in 1987 to 1988, studied two communities to see how family structure influenced fertility decisions. By adding less-structured ethnographic interviews to the highly structured survey, the investigators discovered that a previously unknown factor, the Small Farmers Development Program (SFDP), had a significant influence on fertility decisions (Axinn, Fricke, & Thornton, 1991; Fricke, 2005).

2.2 The event history calendar method is easily adaptable to fit cultural needs. Some tribes in the Chitwan Valley Family Study (CVFS), conducted in Nepal, had no conception of time measurement.
Researchers successfully used local and national events as landmarks to help respondents accurately recall their life course history (Axinn & Pearce, 2006; Axinn, Pearce, & Ghimire, 1999; Belli, 1998).

2.3 To look at trends in household poverty, Krishna (2007) followed seven steps in a Stages of Progress method:

2.3.1 Assembled a "representative community group" (p. 2);
2.3.2 Presented objectives;
2.3.3 Collectively described the construct;
2.3.4 Used current definitions of households as the unit of analysis, inquired about the status of the construct at present and 25 years ago;
2.3.5 Assigned households to categories;
2.3.6 Asked about reasons for descent into poverty among a sample of households within each poverty category (relative to previous and current poverty status); and
2.3.7 Interviewed household members.

2.4 Broom (2005) believes that health research is best conducted using in-depth interviews, rather than being driven by the questionnaire and preconceived notions. He argues that qualitative methods allow for a more thorough analysis and holistic understanding of the patients’ decision-making processes.

2.5 Hulme (2007) describes the use of mixed methods in the context of country case studies.

2.6 The Demographic and Health Surveys (DHS) program conducts research in approximately 75 developing countries across the world (DHS, 2010). The main objectives of the DHS program are "(1) to provide decision makers in the survey countries with data and analyses useful for informed policy choices, (2) to expand the international population and health database, (3) to advance survey methodology, and (4) to develop in participating countries the skills and resources necessary to conduct demographic and health surveys.” Phase II of the DHS introduced a calendar at the end of one of the core questionnaires to clarify dates relating to fertility, contraceptive, postpartum, marriage, migration, and employment history. The researchers found that the calendar provided gains in the quantity and quality of data collected, as well as increasing their analytical potential.

2.7 Hargreaves et al. (2006) used mixed methods to assess the poverty rankings of individual households in eight villages in rural South Africa. The study aimed to identify the number of poor households
and to assess their level of poverty. Working with researchers, community residents drew a map of their village and located each household within its boundaries. Researchers then asked smaller groups of residents to rank pairs of randomly selected households, asking which household in the pair was poorer and which was better-off. Finally, the responses were coded. The authors found strong agreement between the subjects’ coded perceptions of poverty and a household wealth index generated using statistical methods. Howe and McKay used similar methods to study chronic poverty in Rwanda (Howe & McKay, 2007).

2.8 Keller (2007) studied the influence of parents and other socialization factors on human development. Working with young infants and their families in Asia, Latin America, Europe, North America, and Africa, she successfully combined qualitative analyses of interviews and participant observation with quantitative analyses of questionnaires and videotape footage.

2.9 Implementing qualitative methods or ethnorsveys helped University of Chicago researcher Douglas Massey gain greater insight into the reasons behind migration in the U.S. (Massey, 1987).

2.10 By combining data obtained from both statistical and qualitative analyses, Sampson and Laub were able to more accurately explain and identify changes and consistencies in criminological behavior over a convict’s life (Sampson, & Laub, 1998).

2.11 Bamberger, Rao, and Woolcock (Bamberger et al., 2010) suggest returning briefly to the field when writing the quantitative report for more descriptive information or to explore inconsistencies in the data.

3. **Reduce the potential for nonresponse bias** as much as possible.

*Rationale*

Optimal data collections maximize response rates and thereby decrease the potential for nonresponse bias. Nonresponse occurs when survey measures are not obtained from sampled persons, thereby increasing the nonresponse rate. Nonresponse bias occurs when the people who are non-respondents differ from respondents systematically. Although the response rate alone does not predict nonresponse bias (Groves, 2006) a low response rate can be a predictor of the potential for nonresponse bias.

Furthermore, response rates have been dropping differentially across countries due to noncontact and, increasingly, reluctance to participate (de Leeuw & de Heer, 2002).
The coordination of a cross-cultural survey can be centralized or decentralized, with a corresponding focus on either input or output harmonization, as discussed in Study Design and Organizational Structure. These differences in coordination can impact response rates and response bias differentially. For example, in a study using the output harmonization model, where each country uses their own methods and strategies to maximize response rate, nonresponse rates can be calculated and response bias can occur in different ways, whereas in a study using input harmonization, study countries will be limited in adaptation to local contexts, which in turn also impacts response rates and response bias. See Wagner and Stoop (2017) for a more in-depth discussion on nonresponse and nonresponse bias in a cross-national study.

For further discussion of nonresponse bias within the survey quality framework, see Survey Quality.

**Procedural steps**

3.1 Consider the following steps at the community level to reduce non-response before beginning data collection.

3.1.1 Depending upon cultural norms, gain the support of any "gatekeepers" (e.g., community leaders or elders) before attempting to reach individual households.

3.1.2 Make all efforts to raise awareness about the need for high quality surveys and thus the need for people to take part.

3.1.3. Publicize the survey locally to raise awareness and encourage cooperation.
- If most of the population is literate, consider displaying colorful, attractive leaflets on local bulletin boards and in other public spaces.
- Use word-of-mouth channels or local dignitaries (doctors, teachers) as appropriate.

3.2 Send pre-notification letters to sampled households if feasible.

3.2.1 The letter should (1) explain the purpose of the survey, (2) establish the legitimacy of the survey organization and the interviewer, (3) assure confidentiality of answers, (4) notify the household that participation is voluntary, (5) include or announce gifts or incentives and provide information about them, and (6) provide contact information for the organization (see Appendix A for an example of pre-notification letters).

3.2.2 There should be a short timespan between the arrival of the letter and first contact by the interviewer; a time span of several days is ideal. If there is a long delay between the letter
and contact, consider sending a second letter before attempting contact.

3.2.3 Personalize the advance letter with the individual name if possible and appropriate.

3.2.4 Be aware that survey sponsorship may affect both response rates and the accuracy of the actual data. For example, some respondents may fear repercussions if they do not respond to a survey sponsored by a government agency. While this fear may dramatically increase response rates, the quality of the data may be dubious; respondents may feel that their responses are not genuinely confidential if the survey is sponsored by a government agency, and they may not respond openly. In addition, ethical issues arise in such situations (see Ethical Considerations).

3.3 Nonresponse can be assessed and reduced with effective sample management and interviewer management monitoring systems and associated paradata. For an in-depth discussion on the use of responsive designs and paradata to assess nonresponse and nonresponse bias, see Paradata and Other Auxiliary Data.

3.3.1 Study structure and data collection modes may specify what sample management systems are used. In cross-cultural surveys with strong centralized control, a single sample management system may be specified in the contract with local survey organizations.

3.3.2 A good sample management system facilitates evaluating interviewer workload and performance.

3.3.3 Monitor response rates continuously, and produce reports of daily response rates in order to identify data collection procedures that are more or less successful at increasing participation.

3.4 Structure the field staff to aid them in working the sample efficiently and effectively.

3.4.1 Give supervisors the responsibility of assigning sample elements to interviewers and reassigning them when necessary.

3.4.2 Do not allow interviewers to switch sample among themselves without the explicit approval of the supervisor.

3.4.3 Ensure that sample elements are assigned in a way that minimizes travel efforts and costs.

3.4.4 Decide whether interviewers will work alone, in pairs, or in traveling teams (see above and Interviewer Recruitment, Selection, and Training).

3.4.5 Decide whether interviewers and respondents should be matched on some characteristic(s), such as gender or
ethnicity, in order to increase respondent comfort and increase respondent cooperation. If the respondents' characteristics are unknown prior to data collection, develop procedures to make on-the-spot matching possible. For example, to facilitate gender matching, send interviewers into the field in male-female pairs.

3.5 Specify the minimum, the maximum number, and the timing of attempts to contact before the final disposition code is assigned to increase efficiency.

3.5.1 Interviewers should attempt to contact respondents at different blocks of time across the week to increase the probability of reaching the respondent at home.

- The times of day when persons are most likely to be at home vary by culture, location and context. For example, working respondents in the United States are more likely to be reached on evenings and weekends (Groves & Couper, 1998).
- Alternatively, specify the minimum number of times that attempts must be made during daytime hours, during evening hours, and during the weekend (see Kulka & Weeks (1988) for details on call scheduling). Incorporate culture-specific information about likely at-home patterns, such as normal workdays, normal work hours, and holidays. Beware of religious and other cultural norms that restrict interviewing at certain times.

3.6 If appropriate, offer an incentive for participation (Singer, 2002).

3.6.1 Adapt the type and amount of the incentive to local customs. Make yourself familiar with country-specific research on incentives.

3.6.2 According to US- and Canada-based research:

- Present the incentive as a "token of appreciation" for participating in the survey, not as payment for the response.
- Make the token reasonable; it should not be so large that it might raise suspicion about the researcher's or organization's motives or be somehow coercive. It should be generally proportionate to the respondent burden.
- Ideally, provide the incentive prior to the interview. Incentives promised upon the completion of the interview also increase participation, but to a lesser degree (Berk, Mathiowetz, Ward, & White, 1987; Singer, Hoeyyk, Gebler, Raghunathan, & McGonagle, 1999).

3.6.3 Document the use of incentives, including amount and type, time of implementation, and any special strategy, such as
increasing the amount of the incentive in the final weeks of the study.

- According to the existing literature, unconditional prepaid incentives seem to be more effective than conditional incentives paid upon completion of the interview (Koch, Blom, Stoop, & Kappelhof 2009). Thus, eliciting feelings of obligation from the unconditional incentive is more effective than rewarding participation.

- It may be necessary to monitor the extent to which monetary incentives disproportionately encourage the participation of people with low incomes compared to those with high incomes and thereby have an effect on nonresponse bias. If poorer people are usually underrepresented in the achieved sample, monetary incentives might reduce nonresponse bias. If poorer people are already overrepresented, incentives might even increase the nonresponse bias.

- Offering a choice of different types of incentives might attract people from a more diverse background. This can help to reduce an existing nonresponse bias and counteract the potentially selective effect of offering one specific incentive.

- For financial incentives, interviewers may be asked to record that an incentive was given to a respondent; similarly, the respondent may need to sign to indicate receipt.

- In deciding whether to use an incentive, weigh the relative time and cost advantages of using an incentive versus not. Incentives may mean less interviewer time in persuading respondents to participate or less time in refusal conversions. The reduction in interviewer time – and thus costs – must be weighed against the cost of providing incentives.

- See Ethical Considerations for further discussion on the appropriate use of incentives.

3.7 In using a face-to-face or telephone mode, train interviewers to use culturally appropriate reluctance aversion techniques (see Interviewer Recruitment, Selection, and Training).

3.7.1 Social or psychological factors (e.g., reciprocation, consistency, social validation, authority, scarcity, liking) affect respondents' decision in survey participation (Cialdini, 1988). Minimally, train interviewers how to answer anticipated respondent concerns (Groves, Cialdini, & Couper, 1992).

3.7.2 Be aware that local customs and legal limitations may prohibit any attempt to recontact someone who has declined to
participate in the survey. In these cases, using appropriate reluctance aversion techniques becomes especially important.

3.7.3 Make sure that supervisors monitor interviewers closely on respondent reluctance issues.

3.8 If using a face-to-face or telephone mode, consider assigning supervisors or more experienced interviewers to cases where interviewers have been unsuccessful making contact or achieving cooperation.

3.9 Consider switching modes to increase contact and cooperation.

3.9.1 Some studies in the United States employ a mixed mode design in which the least expensive mode is used initially, after which time progressively more expensive modes are implemented in order to reduce nonresponse.

3.9.2 Different modes may produce different survey estimates. These mode-specific differences in measurement might be acceptable to the investigator if nonresponse is sufficiently reduced.

3.9.3 If more than one mode is expected to be used and budget permits, examine possible mode effects prior to the start of data collection.

- Test for mode effects by administering key questions or questionnaire sections to a randomly split sample of respondents similar to the targeted population (e.g., asking the questions on the telephone for one group and in-person for another).
- If it is not possible to test for potential mode effects beforehand, check for differences in responses at the end of data collection.
- Ascertain whether respondents surveyed in each mode produce similar response distributions on key variables before combining their responses for analysis.

3.10 Have interviewers complete a contact attempt record each time they attempt contact, whether or not the attempt is successful (see Appendix B for an example of a contact attempt record).

3.10.1 Use disposition codes to describe the outcome of each contact attempt.

3.10.2 Distinguish (1) completed interviews with eligible persons, (2) non-interviews (eligible persons), (3) non-interviews (unknown if eligible persons), and (4) non-interviews (ineligible persons).

3.11 Assign a final disposition code to each sample element in the gross sample at the end of data collection; include any new sample
elements that may be created or generated during data collection (e.g., for additional family members or through half open intervals).

3.11.1 Provide a clear explanation and training to interviewers before they are allowed to assign final disposition codes.

3.11.2 Take into account that, in some survey organizations, only supervisors can assign final disposition codes.

3.11.3 See Appendices D-G for a description of disposition codes and templates for calculating response rates from the American Association for Public Opinion Research (AAPOR).

3.11.4 See also AAPOR’s Standard Definitions publication (American Association for Public Opinion Research, 2016), which also provides definitions for final sample disposition codes and formulas for calculating response, refusal, and other rates. See also AAPOR’s Response Rate Calculator (available for download at: http://www.aapor.org/Standards-Ethics/Standard-Definitions-(1).aspx?utm_source=AAPOR-Informz&utm_medium=email&utm_campaign=default).

3.11.5 Note that the list of disposition codes may need to be modified for the local situation and additional codes may need to be defined to account for local conditions.

3.12 Minimize the effects of nonresponse bias on analyses as much as possible.

3.12.1 Nonresponse bias is a function of both the response rate and the difference between respondents and nonrespondents on a particular statistic (Groves & Couper, 1998). Because nonresponse bias is statistic-specific, response rates alone do not indicate nonresponse bias. Therefore, estimate the effect of nonresponse bias on key survey estimates, if possible (see Guideline 7 below).

3.12.2 If possible, use weighting and imputation (Groves, Dillman, Eltinge, & Little, 2002) (see Data Processing and Statistical Adjustment).

Lessons learned

3.1 Differences in response rates cross-nationally can be due to many factors, including differing judgments of interviewers and other local survey staff about the efficacy and subsequent application of particular survey research techniques and protocols. A review of response rates from the 1995 round of the International Social Survey Programme (ISSP) found significant differences in response rates, with at least some of the difference likely attributable to mode (face-to-face vs. mail). Even for countries with roughly comparable response rates, sources of nonresponse differed, with noncontact...
contributing substantially to nonresponse in Japan, and refusal contributing to nonresponse in Russia (Couper & de Leeuw, 2003).

3.2 Response rates are not necessarily good indicators of nonresponse bias, but nevertheless tend to be used as a proxy for bias. In a health study of the elderly in Scotland, healthy individuals were more likely to participate than unhealthy individuals. Because of this difference between the respondents and nonrespondents, the estimate of health was biased even though response rates reached 82% overall (Cohen & Duffy, 2002).

3.3 While the literature has clearly established the positive effects of prepaid and cash incentives upon response in minority countries (Berk et al., 1987; Singer et al., 1999), it is possible that incentives may affect the propensity to respond differently among a population with high rates of poverty. For example, offering a choice of incentives may be more effective at increasing response rates than simply offering a prepaid incentive. Furthermore, in areas with rampant inflation, the value of cash incentives may decrease dramatically within a short period of time.

3.4 The same incentive may affect response rates differently across countries or cultures. In the German General Social Survey (ALLBUS), the same incentive (€10) was offered to all respondents. The authors examined cooperation rates for Moroccan and Turkish immigrants. The authors found that the incentive affected cooperation differently by ethnicity and gender: cooperation rates increased as a result of the incentive for Moroccan women, but did not increase for Moroccan men, Turkish men, or Turkish women (van den Brakel, Vis-Visschers, & Schmeets, 2006).

3.5 The mechanism of incentive efficacy will differ across mode. In telephone surveys, incentives are often sent to the respondent in an advance letter prior to contact, to encourage cooperation. In mail surveys, the incentive may be sent either in advance or along with the mailed questionnaire. And, in face-to-face interviews, the respondent generally receives the incentive at the conclusion of the interview, meaning that the actual transfer of the incentive, and therefore its effect on response rate, can differ across mode, leading to further differentiation in response rates cross-nationally if different countries use different modes in a cross-national survey.
3.6 Use caution when choosing to give monetary awards to study participants. Keller studied the influence of parents and other socialization factors on human development in Asia, Latin America, Europe, North America, and Africa. Respondents received a cash incentive. Keller experienced some hostility from families that were not selected for the study (and, thus, not given any monetary rewards) because they did have young children (Keller, 2007).

3.7 Some studies vary incentive use within a country; for example, offering incentives only to respondents in urban areas, where response rates are typically lower; or offering incentives only in cases of refusal, in an attempt to gain cooperation. If considering this approach, be aware of any concerns that might arise from ethics review boards.

3.8 Countries have different incentive norms.
   3.8.1 For example, in a recent study conducted in Nepal and the United States, respondents in Nepal were highly cooperative and were offered no financial incentive. In the U.S., however, potential respondents were not as cooperative or easy to contact, and incentives were required (Axinn, Chardoul, Ghimire, Gatny, & Barber, 2008).
   3.8.2 Some 3MC surveys (e.g., the European Social Survey and the Living Standard Measurement Study Survey (Living Standard Measurement Study Survey, 1996) allow each participating country to decide whether or not to offer incentives.
   3.8.3 If incentives are offered, the type may vary from one country to another. For example, the Survey of Health, Ageing and Retirement in Europe (SHARE) offers various incentives, depending on the country's culture. Incentives for the World Mental Health Survey (Keller, 2007) vary across participating countries, including but not limited to, cash (in the Ukraine and United States), an alarm clock (in Columbia), and a bath towel (in Nigeria); no respondent incentives are offered in Mexico, South Africa, Belgium, Germany, Israel, Japan, or China. In the Netherlands, flowers are a customary gift to the hostess when visiting for dinner, and therefore flowers are an effective incentive in the Netherlands.

3.9 Similarly, many cross-cultural surveys (e.g., the European Social Survey, the Living Standard Measurement Study...
Survey (Living Standard Measurement Study Survey, 1996), and the World Mental Health Survey (Kessler, Ustun, & World Health Organization, 2008)) allow participating countries to vary in their use of advance letters and follow-up letters. In the Survey of Health, Ageing and Retirement in Europe (SHARE), advance letters are mailed to each household in the gross sample and follow-up letters are used with reluctant respondents.

3.10 In an experimental design in the U.S., researchers investigated the use of a novel incentive they termed "reciprocity by proxy", wherein respondents were invited to participate in a program with the promise that their participation would result in a gift to a third party, such as a charity. Researchers found that reciprocity by proxy increased participation more than either incentive by proxy or no incentive. However, researchers caution that this approach can backfire if the target audience does not support the beneficiary of the gift (Goldstein, Griskevicius, & Cialdini, 2012). To mitigate this risk, researchers can offer to make a contribution to a charity of the respondent’s choosing.

3.11 An effective sample management system can clarify the causes of nonresponse. When the Amenities and Services Utilization Survey (AVO) was conducted in the Netherlands in 1995, interviewers were not asked to record detailed disposition codes for each call. As a result, refusals could not be distinguished from noncontacts. When the study was repeated in 1999, detailed disposition codes were collected. Researchers were then able to see that, after three unsuccessful contact attempts, refusal was the more probable explanation (Stoop, 2005).

3.12 Not all survey organizations will be familiar with sample management practices. Allow some time in training for interviewers to become familiar with the sample management system (see Interviewer Recruitment, Selection, and Training) and check completed forms.

4. **Time data collection activities appropriately.**

*Rationale*

A specific survey estimate of interest may determine the timing of data collection activities; for example, a survey about voting behavior will necessarily be timed to occur around an election. Data collection activities
may be hampered by inappropriate timing. Face-to-face data collection, for example, may be impossible during a monsoon season, an earthquake or a regional conflict.

The guideline assumes that a specific start time and end time to data collection exists; this guideline does not address issues in continuous data collection.

**Procedural steps**

4.1 Based upon feasibility studies (see Guideline 1 above), evaluate environmental, political, and cultural considerations which might affect the timing of data collection. These could include:
   4.1.1 Extreme weather patterns or natural disasters.
   4.1.2 War, civil war, military rule, militia rule, or the possibility of hostage taking.
   4.1.3 Religious and secular holidays or migratory patterns of nomadic people. For example, Independence days (e.g., Bastille Day in France), New Year’s Day in China, summer Christmas holiday in Australia and New Zealand, and vacations in July and August in Europe would not be a good time.

4.2 Establish a specific start and end date for data collection.
   4.2.1 Keep a concurrent fielding period across countries. This would guarantee the cross-national comparability. For example, the ESS requires interviewers across participating countries in Europe to collect data within a four-month period from September to December of the survey year (Koch et al., 2009).
   4.2.2 If the 3MC project includes countries located in both the northern and southern hemispheres, where summer and winter are in opposition, consider what field period is most feasible for all countries.
   4.2.3 Because unexpected events can interfere with data collection activities, remain somewhat flexible to allow for unexpected events. Include details about any deviations from the anticipated schedule in the study documentation.

**Lessons learned**

4.1 Coordination of data collection activities across countries or cultures can be difficult or even impossible. The Afrobarometer measures public opinion in a subset of sub-Saharan African countries. The coordinators for the Afrobarometer note that data collection is especially difficult during national election or referendum campaigns,
rainy seasons, times of famine, and national or religious holidays. Since such events vary across countries and cultures, fieldwork activities are spread over a full year (Afrobarometer Survey, 2010).

4.2 Timing of data collection activities may be related to the topic of the survey or statistics of interest. The Comparative Study of Election Systems (CSES), for example, studies elections around the world and therefore must time data collection activities according to local election cycles (Howell, 2010).

4.3 The response rate for the Asian Barometer survey in Japan in 2003 was 71%. In 2007, the response rate dropped to 34.3%. One possible reason for the sharp drop in response rates in 2007 is that, in 2006, the law no longer allowed commercial surveys to use voter lists or resident registries. As a result, many people mistakenly believed that the new regulation also applied to academic research (Ikeda, Yamada, Taniguchi, Nishizawa, & Steel, 2007).

4.4 Data collection in Germany for the first European Social Survey had to be delayed due to general elections held in that autumn.

4.5 In some settings, electrical availability is dictated by the calendar and should be evaluated prior to data collection. For example, Nepal relies primarily on hydropower, and so electricity shortages increase significantly in most areas of the country during the dry season between February and April, with some areas without electricity for more than 14 hours per day. Recharging equipment in these sorts of environments can be a major impediment (Paudel, Ahmed, Pradhan, & Dangol, 2013).

5. Institute and follow appropriate quality control measures.

Rationale

If errors are caught early, they can be corrected while the study is still in the field. Improvement made during data collection may introduce some measure of inconsistency in the data, however. This trade-off should be considered before any action is taken (Groves, 2006). See also Survey Quality for a discussion of the quality control framework and Paradata and Other Auxiliary Data for a detailed discussion on using paradata in quality control and survey error reduction.

Procedural steps

5.1 Evaluate the effectiveness of data collection protocols regularly. Include:
5.1.1 Sample management systems.
5.1.2 Contact protocols.
5.1.3 Reluctance aversion protocols.

5.2 With real-time or daily data transmission, quality control routines and error detection can be implemented more efficiently
5.2.1 The use of technology for data collection allows for collecting and analyzing paradata (such as keystrokes and time-stamps, GPS coordinates) for monitoring interviewer behavior (if an interviewer-administered mode is used). This allows for early detection of interviewer deviation from interviewing protocol, for early intervention, and better data quality. [See Lessons Learned below as well as Paradata and Other Auxiliary Data]. Moreover, post-survey processing time is greatly reduced.
5.2.2 If an interviewer-administered mode is used, observe the interviewers throughout data collection (Lavrakas, 1993); monitor them more frequently early in the study, less frequently as the study continues.

5.3 If an interviewer-administered mode is used, review a random sample of coversheets on an ongoing basis to ensure that correct eligibility and respondent selection procedures are being followed.

5.4 If an interviewer-administered mode is used, provide interviewers with feedback, both individually and as a group (Couper, Holland, & Groves, 1992; Lavrakas, 1993).
5.4.1 Provide immediate, individual feedback if there has been a critical error.
5.4.2 Provide routine, individual feedback for self-improvement.
5.4.3 Offer group feedback to focus efforts on improving the process.
5.4.4 Continually evaluate the following with respect to interviewers (Biemer & Lyberg, 2003):
  - Knowledge of the study objectives.
  - Administration of the survey introduction.
  - Administration of household enumeration and respondent selection procedures.
  - Reluctance aversion efforts.
  - Contact efforts.
  - Rapport with the respondent (e.g., having a professional, confident manner).
  - Standardized interviewing techniques (e.g., reading questions as worded, probing, and clarifying).
  - Data entry procedures.
  - Administrative tasks (e.g., submitting timesheets in a timely fashion).
- Ability to meet production goals and maintain productivity.
- Administration of specialized study-specific procedures (e.g., procedures for taking physical measurements and administering tests of physical performance or cognitive ability).

5.5 Whenever possible, recontact or reinterview approximately 10-15% of each interviewer’s completed cases, selected at random (American Association for Public Opinion Research, 2003; Office of Management and Budget, 2006).

5.5.1 If recontacting the respondent, verify that the interview took place, inquire if interviewer acted professionally, and ask factual questions (e.g., mode of data collection, interview length, incentive, household composition, and key survey topics (American Association for Public Opinion Research, 2003)).

5.5.2 If reinterviewing the respondent, ask a sample of factual questions that do not have heavily skewed response distributions, were not skipped by many respondents, are scattered throughout the questionnaire, and have answers which are unlikely to have changed between the time of the interview and the verification check (Forsman & Schreiner, 1991; United Nations, 2005).

5.5.3 Conduct reinterviews within a time period that is not so long that respondents will have forgotten about the survey or so short that respondents will remember all the details of the survey (Forsman & Schreiner, 1991).

5.5.4 Make sure recontacts and reinterviews are made with the original respondent and that questions refer to the same time period as was asked about in the original interview (Forsman & Schreiner, 1991).

5.5.5 In some countries, it is not possible to perform recontacts or reinterviews due to laws and/or local customs. Document such instances.

5.6 If feasible, audio record face-to-face interviews for review.

5.6.1 Determine whether cultural norms permit taping.

5.6.2 Inform respondents that they may be recorded for quality purposes and allow respondents to refuse to be recorded.

5.6.3 Store any tapes safely and securely (see Ethical Considerations).

5.7 Identify potential interviewer falsification.

5.7.1 Implement silent monitoring in centralized facilities, use audio-recordings and recontacts in field studies, and
analyze outliers in the data to detect falsification (American Association for Public Opinion Research, 2003).

5.7.2 Check responses to stem questions for each interviewer. Questions that have a stem-branch structure—in which specific responses to "stem" questions require the interviewer to ask a number of "branch" questions—can be at increased risk for falsification. If a particular interviewer has recorded responses to stem questions that consistently preclude the interviewer from asking the branch questions, the interviewer may be falsifying data.

5.7.3 Examine paradata, such as keystroke data and time stamps, by interviewer to identify potential falsification.

5.7.4 Examine survey data for any duplicate cases, which can indicate falsification as well as data processing error.

5.7.5 If falsification of data is suspected, contact the respondents involved over the telephone (Forsman & Schreiner, 1991). If respondents cannot be reached via telephone, send out a brief mail questionnaire with a prepaid return envelope (Biemer & Lyberg, 2003).

5.7.6 If falsification of data is suspected, investigate the interviewer's other work and remove the interviewer from all data collection activities until the issues have been resolved (American Association for Public Opinion Research, 2003).

5.7.7 If irregularities or falsified data are discovered, redo the interviewer's cases and delete all of his or her recorded data (American Association for Public Opinion Research, 2003; Biemer & Lyberg, 2003).

5.8 For approximately 5% of each interviewer's finalized non-interviews, perform random checks with households to verify that ineligibility, refusal, or other status was correctly assigned. Checks may be done by telephone, in person, or by mail, as needed.

5.9 If physical measurements are being taken:

5.9.1 Periodically retest the interviewers on the use of these instruments.

5.9.2 Select equipment that can withstand the local conditions (heat, cold, altitude, etc.).

5.9.3 Document the technical specifications of the equipment chosen.

5.9.4. Re-calibrate equipment as needed throughout data collection.

5.10 If the survey is being conducted in a centralized telephone facility, follow established monitoring procedures (Couper et al., 1992).
5.10.1 Monitor in relatively short (e.g., one-hour) shifts; this is cost-effective and reduces supervisor fatigue.

5.10.2 Use probability sampling to ensure that the number of interviews monitored is proportional to the number of interviewers working each hour (see Sample Design).

5.10.3 Monitor new interviewers at a higher rate than experienced interviewers.

5.10.4 Select from eligible cases in which the phone is still ringing so that the supervisor is not forced to wait for new interviews to begin in order to start monitoring.

5.11 Monitor quality indicators consistently throughout the field period; use an electronic system or note them in a daily log book (United Nations, 2005). Include the following:

5.11.1 Distributions of key variables.

5.11.2 Hours per interview per interviewer, for the study as a whole, and by respondent groups of interest.

5.11.3 Number of respondents approached, interviews completed, incomplete interviews, and contact attempts.

5.11.4 Response, refusal, and non-contact rates (United Nations, 2005) (see Data Processing and Statistical Adjustment).

5.11.5 Outcomes of all contacts and review of disposition code assignment.

5.12 Create statistical process control charts (SPCs) to provide timely information on key aspects of the data collection process (Mudryk, Burgess, & Xiao, 1996).

5.12.1 Use the charts to detect observations that are not within predetermined limits (often between one and three standard deviations of the mean).

- A common use of SPCs in survey organizations is to assess nonresponse reduction methods over the field period. Using these charts, the impact of interviewer effort on response rates can be easily assessed (see case studies in Survey Quality for additional discussion of SPCs).

5.12.2 Give extreme observations additional attention and try to determine the root cause.

5.12.3 Refer to the charts when deciding whether to release additional sample elements for interviewers to attempt to contact, further monitor interviewers, and offer additional training sessions.

5.13 Set contact limitations, determining:

5.13.1 The point at which additional attempts to contact a sample element are inefficient.
5.13.2 Whether respondents cooperating after a certain number of contact attempts are significantly different from others on key indicators.

Lessons learned

5.1 **Process** and **progress indicators** are often interdependent. Therefore, improving one process or progress indicator may negatively affect another, particularly in the context of attempts to achieve cross-national comparability. For example, the pursuit of higher response rates can actually increase **nonresponse bias** if the techniques used to obtain the higher response rates are more acceptable and effective in some cultures than in others ([Groves, 2006; Harkness, 1999](#)).

5.2 In Round 4 of the **Afrobarometer Survey (2010)**, teams of four interviewers travel together to the field under the leadership of a field supervisor who has at least an undergraduate degree and experience in collecting data and managing field work teams or no degree but extensive experience. It is the supervisor’s job to ensure quality control of survey returns on a daily basis. Interviewers are monitored at all stages and debriefed daily immediately after interviews. Completed questionnaires are checked for missing data and inconsistencies. Each field supervisor maintains a daily written log of observations on sampling and interviewing conditions and political and economic features of the area and makes daily telephone report to headquarters. A fieldwork debriefing is held after all returns have been submitted. Sampling back-checks are routinely conducted to ensure that the respondent selection is correctly done. The field supervisor also verifies basic information (e.g., respondent age and level of formal education).

5.3 The **Asian Barometer Survey (2010)** required all interview teams to travel together under the supervision of a field supervisor and to have a debriefing meeting each evening. Supervisors randomly checked with respondents to make sure the interviews were done properly.

5.4 In Round 5 of the ESS ([ESS, 2010](#)), quality control back-checks were performed for at least 10% of respondents and 5% of the non-respondents either in person, by telephone, or by mail. For the respondents, a short interview was conducted to confirm the interview, whether showcards were used, the approximate length of the interview, etc.

5.5 In the Living Standard Measurement Study ([LSMS, 1996](#)), each field supervisor oversees two interviewers. Each week the field supervisor
observes and evaluates one interview per interviewer and documents the process for submission to the national office. Data collection is broken into two rounds; the first half of the questionnaire is completed in round one and then checked for accuracy before the second half of questionnaire is completed in round two. After the second round, only data entry errors are corrected. Check-up interviews are routinely performed in 15% to 25% of the households.

5.6 The Survey of Health, Aging and Retirement in Europe (SHARE) requires all survey agencies to use an electronic sample management system (SMS). All but three participating countries (France, the Netherlands, and Switzerland) use a "Case Management System" (CMS), developed by CentERdata. This system monitors the survey progress in real-time, including screening for eligible respondents, recording contact attempts, ensuring the correct implementation of contact and follow-up strategies, and managing refusal conversion strategies. Bi-weekly reports are generated for the coordinating team.

5.7 The recommended supervisor-to-interviewer ratio in the World Mental Health Survey is 1 for every 8 to 10 experienced interviewers, with those countries using a pencil-and-paper mode having a higher ratio than those conducting computer-assisted surveys. Supervision consists of direct observation and/or audio recording of part or all of the interview for 5% to 10% of each interviewer's work. Supervisors randomly select 10% of interviewed households, confirm the household listings and selection procedure, and repeat some of the questions. Open-ended responses and other quality control checks are reviewed on a daily basis by supervisors, and interviewers recontact respondents to obtain missing data (Kessler et al., 2004; Kessler et al., 2008).

5.8 Data falsification can be difficult to detect and there is no one identification strategy. Kuriakose and Robbins (2015) suggest researchers set a benchmark (in this example, 85%), wherein any two cases where at least 85% of responses are duplicate to be suspicious. However, this strategy has been argued to produce a large number of false positives (Bohannon, 2016, Simmons, Mercer, Schwarzer, & Kennedy, 2016), and researchers argue that each survey has unique parameters that researchers should account for when analyzing data for potential falsification.

5.9 In surveys conducted at the Allensbach Institute in Germany, researchers have used two different methods to mitigate interviewer falsification in lieu of recording respondent contact information and performing post-survey verification (Smith, 2011a). In the first method,
researchers included a factual question in the survey that asked about a little-known fact that would be unanswerable to most respondents. Later in the survey, a second item provided the information that would answer the earlier factual question. In a valid interview, respondents would not be able to go back in the questionnaire to use this information to answer the first question correctly; therefore, it was expected that the vast majority of respondents would provide the wrong answer to the first question. However, an interviewer falsifying responses could potentially use the information to correctly answer the first item. Researchers could then identify any interviewer whose respondents had accurate responses for the first survey question and investigate his or her other completed interviews for a pattern indicating possible falsification. A second technique used by the researchers in Allensbach was to have respondents write responses to open-ended questions. The handwriting could then be examined to see if the interviewer was completing the interviews him- or herself.

6. Document data collection activities.

Rationale

The documentation of data collection procedures is an essential part of the data collection process. Process documentation is necessary for timely intervention. In addition, by understanding what was done in the field, the data are more easily interpreted and understood.

Procedural steps

6.1 Document the following (see Appendix C):

6.1.1 A summary of feedback from the feasibility studies.
6.1.2 The interview or data collection process.
6.1.3 A description of the mode(s) used.
6.1.4 A description of the mode-specific protocols.
6.1.5 A description of the sample management system.
6.1.6 A description of any paradata collected.
6.1.7 Special approaches to reduce nonresponse, including any incentives and nonresponse follow-up.
6.1.8 Outcome rates by key respondent groups, including response, refusal, noncontact, and other nonresponse rates.
6.1.9 Structure of the field staff (e.g., size of interviewer groups and supervisor/interviewer ratio).
6.1.10 Timing of the fieldwork for each country or cultural group.
6.1.11 A description of quality control procedures and protocols, including:
   - Interviewer monitoring procedures.
Outcomes of interviewer monitoring, such as hours per interview and any falsification rates.

6.1.12 Any validation study descriptions and outcomes (see Guideline 7 below)

7. When possible, conduct validation studies to estimate bias.

Rationale

As noted in Guideline 3 above, response rates alone are not good indicators of nonresponse bias; understanding nonresponse bias and making subsequent post-survey adjustments require information about the nonrespondents. Similarly, measurement error bias can only be estimated when "true" values for survey variables are known or can be modeled (i.e., using latent class analysis). Validation studies can increase confidence in results, assist with post-survey adjustments (see Data Processing and Statistical Adjustment), and address potential criticisms of the study. However, while the interpretation of survey estimates can benefit greatly from validation studies, conducting them may be difficult and prohibitively expensive.

Survey methodological experiments are designed up front and the outcomes are carefully documented. While these experiments may or may not directly benefit a given study, they are extremely important for the development and building of a body of knowledge in cross-national survey methodology, on which future studies will be able to draw.

Procedural steps

7.1 Collect data on nonrespondents, if possible, to estimate nonresponse bias (Groves, 2006).

7.1.1 One approach is to study sample elements that initially refused to be interviewed.

- Draw a random sample of such initial nonrespondents and attempt to interview them under a modified design protocol (e.g., increased incentives or a shorter interview).
- This approach assumes that people who were initially reluctant to participate are identical to nonrespondents on key variables; this may or may not be a valid assumption (Lin & Schaeffer, 1995).
- Document the data collection procedures, including the proportion of initial nonrespondents included in the validation study, mode of administration, and any additional incentive (Groves & Heeringa, 2006).
7.1.2 A second approach is to compare respondents and nonrespondents on statistics of interest using information contained in external records (e.g., population register data).
- Complete external records for all sample elements may be difficult to find, inaccurate, or outdated.
- These benchmark data are rarely available for statistics of interest.

7.1.3 A third approach is to calculate response rates within subgroups (e.g., racial, ethnic, or gender groups).
- This approach assumes that subgroup membership is related to the propensity to respond, and assumes that biases in demographic variables are informative of biases in substantive variables.

7.1.4 A fourth approach is to compare estimates to similar estimates generated from outside surveys.
- While estimates similar to estimates from these benchmark surveys can increase credibility, the key survey variables may not exist in the benchmark survey. Furthermore, coverage, nonresponse, and measurement error differences in the benchmark survey are largely unknown.

7.1.5 A fifth approach is to examine the effect of post-survey adjustments on the estimates by comparing unadjusted and adjusted values.
- The use of this approach strongly assumes that the models used to adjust for nonresponse fully capture the nonresponse mechanisms at work. While some amount of nonresponse bias may be controlled using these adjustments, they will rarely—if ever—fully control nonresponse bias.
- See Data Processing and Statistical Adjustment for more information on post-survey adjustments for nonresponse.

7.2 Use methodological studies to assess measurement error.

7.2.1 One approach is to use cognitive laboratory techniques, such as cognitive interviews, vignettes, response latency, and behavior coding (see Pretesting), to assess potential measurement error.
- This approach assumes that laboratory measurements are comparable with those obtained in the survey.
- Many laboratory experiments do not use probability-based samples; therefore, errors detected in the self-selected laboratory sample may not be representative of errors in the target population.

7.2.2 Another approach is to check outside records for the true value, or a proxy of the true value, of the measure.
The researcher must have access to the outside records.
- This approach assumes that the outside records are complete and error-free.
- It may be difficult to match the respondent to the outside record.
- Document record collection procedures, including a description of the records and their quality.

7.2.3 A third approach is to embed a randomized experiment within the survey to assess differences in survey estimates among different measurement conditions. In this situation, respondents should be randomly assigned to the experimental conditions (e.g., interview mode).

7.3 Consider using other methods of assessing measurement error.

7.3.1 Reinterview respondents. Reinterviews are especially useful in determining interviewer falsification (Forsman & Schreiner, 1991) but may also help assess other forms of measurement error (see Biemer, 2004; Biemer & Forsman, 1992) for details on estimating simple response variance or bias).

7.3.2 Document all aspects of the reinterview procedure, including:
- The respondents who were eligible for the reinterview component of this study (e.g., random 10% of respondents), as well as the total number of respondents selected and how many completed the reinterview.
- The questionnaire used in the reinterview.
- The mode of administration of the reinterview.
- The interviewers who administered the reinterview (e.g., any project interviewing staff, specially designated interviewers, supervisory staff, clinicians, self-administered, etc.).
- The time interval between administration of the main interview and the reinterview (e.g., reinterviews were conducted 1-2 weeks after the main study interview).

7.3.3 Collect paradata that may be correlated with measurement error (e.g., number of keystrokes, length of interview).

7.3.4 Use interpenetration to estimate correlated response variance due to interviewers.

Lessons learned

7.1 Supplemental studies can be difficult and expensive to implement, but they are useful for validating survey results. For example, a study of discharged patients at a French hospital found no difference in patient satisfaction ratings between early and late respondents. The authors interpreted this finding to indicate that there was little evidence of nonresponse bias in their estimates of patient
satisfaction. However, it is unclear if the differences in estimates were due to nonresponse bias or to measurement error (Gasquet, Falissard, & Ravaud, 2001).

7.2 Try to use resources to gain knowledge on bias in an efficient way. Validation studies are expensive but come late. Therefore, one should first strive for more preventive measures that hopefully make processes almost error-free. Then paradata should be collected and analyzed so that processes can improve and display a decreased variability. Finally, some small-scale validation studies, rather than large ones, should be conducted, and used as input to more long-term improvements of processes and methods. The optimal allocation between the three is unknown but the general preferred allocation is evident, namely prevention first, then process adjustments via paradata, and lastly small validation studies.
Appendix A

Example of pre-notification letter from ESS 2008 (*Forsman & Schreiner, 1991*).

Autumn 2008

[Dear ....]

European Social Survey 2008

[You have/your address/household has] been selected to take part in an international study on what people think about various important issues affecting [country]. The study is being carried out simultaneously in 25 countries across Europe and will help to find out how much or how little people in different countries share the same views and beliefs. It is being paid for from both [country] and European sources.

The questionnaire covers a wide range of topics and no special knowledge is needed to answer any questions. Your [name/household/address] has been selected from [sampling frame] by scientific methods to ensure that we get a representative picture of people in [country]. We cannot therefore substitute any [name/household/address]. All information you provide will be treated in strict confidence and will never be linked to your name or address.

Most people taking part in the study find it an interesting and enjoyable experience, and we hope that you will too. An interviewer will [visit] you shortly to explain more about the study and, if you agree, will arrange a suitable time for the actual interview, which will be carried out in person. Interviews normally take just under an hour. We certainly hope we can rely on your co-operation.

Meanwhile, if you wish to have any further information about the study, please feel free to contact me on the number above.

Thank you in advance for your help,

xxxx xxxxxx

The interviewer who will be contacting you is:..........................................................................................
Appendix B

Contact attempt record (example from the University of Michigan’s Institute for Social Research)

<table>
<thead>
<tr>
<th></th>
<th>CALL #1</th>
<th>CALL #2</th>
<th>CALL #3</th>
<th>CALL #4</th>
</tr>
</thead>
<tbody>
<tr>
<td>DATE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAY OF WEEK</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EXACT TIME BEGAN</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IVER ID</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CONTACT WITH</td>
<td>R/INF/NO ONE</td>
<td>R/INF/NO ONE</td>
<td>R/INF/NO ONE</td>
<td>R/INF/NO ONE</td>
</tr>
<tr>
<td>MODE OF CONTACT</td>
<td>PERSONAL/TEL</td>
<td>PERSONAL/TEL</td>
<td>PERSONAL/TEL</td>
<td>PERSONAL/TEL</td>
</tr>
<tr>
<td>TELEPHONE NUMBER IF OBTAINED</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HU LISTING OBTAINED</td>
<td>YES/NO</td>
<td>YES/NO</td>
<td>YES/NO</td>
<td>YES/NO</td>
</tr>
<tr>
<td>DETAILED DESCRIPTION OF CONTACT OR CONTACT ATTEMPT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DISPOSITION CODE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

R = Respondent    HU = Housing Unit
Inf = Informant    Listing = enumeration
Appendix C

Documentation

<table>
<thead>
<tr>
<th>Details</th>
<th>Examples from SMDS\textsuperscript{1}</th>
<th>Examples from ISSP\textsuperscript{2}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data collection organizations</td>
<td>The number of organizations</td>
<td>How many organizations conducted data collection for this study in your country? If your agency/organization contracted with another organization which provided data collection services, please include that here.</td>
</tr>
<tr>
<td>Contact information</td>
<td>Type of organizations (e.g., government agency, private research company)</td>
<td></td>
</tr>
<tr>
<td>Data collection methods</td>
<td>The number of separate survey data collection efforts</td>
<td>How were the face-to-face interviews administered in this study? Please check all that apply.</td>
</tr>
<tr>
<td>A brief title of each survey data collection efforts</td>
<td>Delivery of sample to interviewers (e.g., computerized sample management system)</td>
<td>- Computer-assisted personal interviewing (CAPI)</td>
</tr>
<tr>
<td>Mode of data collection</td>
<td>Screening/respondent selection procedure</td>
<td>- Paper and pencil interviewing (PAPI)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Other, specify: _____________________</td>
</tr>
<tr>
<td>Techniques used to maximize response rate</td>
<td>Pre-notification strategies</td>
<td>Which, if any, of the following pre-notification strategies were used for the face-to-face interviews that were conducted in this study? Please check all that apply.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Advance letter sent prior to initial visit</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Email message sent prior to initial visit</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Telephone call made prior to initial visit</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Announcement in local newspaper, radio, or television</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Other, specify: _____________________</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- None of the above</td>
</tr>
<tr>
<td>Use of incentives</td>
<td>Specific incentive offers made to a particular group of sample members</td>
<td>How many different respondent incentives were initially used for the face-to-face interviews that were conducted in this study? For example, if half of the respondents were randomly assigned to receive â‚¬15 and the other half received a gift basket, this should be recorded as two incentives; or if â‚¬10 was included in the advance letter and â‚¬20 was promised upon completion of the interview, this should also be counted as two incentives.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Details

<table>
<thead>
<tr>
<th></th>
<th>Examples from SMDS&lt;sup&gt;1&lt;/sup&gt;</th>
<th>Examples from ISSP&lt;sup&gt;2&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>___ different incentives used {1-10}</td>
<td></td>
</tr>
<tr>
<td>- Refusal conversion protocols</td>
<td>Which, if any, of the following (additional) techniques were used to maximize response rate for the face-to-face interviews that were conducted in this study. Please check all that apply.</td>
<td></td>
</tr>
<tr>
<td>- Interviewer incentives/ bonuses</td>
<td>- Special refusal aversion or refusal conversion training sessions for interviewers</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Specially designated interviewers for reluctant cases (e.g., flying experienced interviewers in to attempt difficult cases)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Persuasion letters sent to reluctant sample members</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Increased or additional respondent incentives implemented after the start of data collection</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Interviewer incentives/bonuses</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- None of the above</td>
<td></td>
</tr>
</tbody>
</table>

### Contact protocols

|  | Minimum number of contacts (on weekday, in the evening, on weekends) before a case is finalized |  |
|  | Maximum number of contacts after a case would be finalized |  |
|  | Methods used to attempt to reach sample members |  |

Was there a minimum number of attempts required before a sample case was finalized (i.e., no more attempts were made on the case) for face-to-face data collection in this study?

- Yes
- No

Were interviewers required to make a certain number of calls/visits before they stopped approaching an address or household?

Minimum number of calls/visits required - please write in number _____

No minimum call requirement ____

### Eligibility screening

|  | Minimum number of attempts for screening (on weekday, in the evening, on weekends) before the case is finalized |  |
|  | Maximum number of attempts for screening after a case would be finalized |  |
|  | Methods for refusal conversion for eligibility screening |  |
|  | Methods to reach sample members for the screening |  |
|  | Any additional techniques that were used to increase response rate for the screening to determine eligibility |  |

What was the minimum number of attempts required before a case was finalized (i.e., no more attempts were made on the case)? If the mode of contact was not specified, please only provide the total number of attempts below.

___ minimum face-to-face attempts {ALLOW VALUE,1-40}

___ minimum telephone attempts {ALLOW VALUE,1-40}

___ total minimum attempts (face-to-face and telephone) {ALLOW VALUE,1-40}

Was substitution or replacement permitted at any stage of your selection process or during fieldwork?

Yes? Question 67
No? Question 68

### Use of special test or data collection besides survey

Besides the survey questions, did this study involve any of the following? Please check all that apply.

- Tests of physical performance (e.g., walking speed, grip strength)

---

**Data Collection: General Considerations**

Revised August 2016
<table>
<thead>
<tr>
<th>Details</th>
<th>Examples from SMDS&lt;sup&gt;1&lt;/sup&gt;</th>
<th>Examples from ISSP&lt;sup&gt;2&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>questions</td>
<td>- Tests of cognitive ability (e.g., memory tasks, word recognition) - Physical measurements (e.g., height, weight, blood pressure) - Collection of biological specimens (e.g., blood, saliva, urine) - Collection of environmental specimens (e.g., soil, dust) - Procurement of respondent permission to access and link respondent data from other sources (e.g., government records, medical records, employment records) - Other, specify: {TEXT BOX} - None of the above</td>
<td></td>
</tr>
</tbody>
</table>

| Locating sample members | Tracking procedures  
• leader/coordinator of tracking  
• tracking manual/tracking team  
• training  
• between wave tracking efforts  
• steps/options used in tracking (relatives, friends, neighbors, and employers) | Were any tracking activities carried out to locate sample members in this study? Please check all that apply.  
- Yes  
- No {SKIP TO DC223} |

| Quality control | Supervision  
• Back-checking | Were any interviews back-checked (e.g. supervisor checks later whether interview conducted)?  
Yes - please write in approximate proportion %  
____  
No ____ |

<sup>1</sup> Survey Metadata Documentation System (SMDS): a standardized web-based documentation tool which was developed by the University of Michigan's Institute for Social Research and Gesis-ZUMA.

<sup>2</sup> International Social Survey Programme (ISSP): see Scholz & Heller (2009), for details.
Appendix D

Disposition codes (American Association for Public Opinion Research, 2016)

- The coordinating center should provide a list of specific disposition codes with a clear description of how to code all sample elements during (temporary disposition) and at the close of (final disposition codes) the field period.

- Generally, disposition codes identify sample elements as (complete or partial) interviews or non-interviews.
  - The coordinating center should set the criteria for determining whether interviews are classified as complete or partial.
  - Non-interviews are grouped by whether the respondent is eligible, unknown eligible, or ineligible to participate in the study.

- Disposition codes are mutually exclusive. While sample elements may be assigned different temporary disposition codes throughout the field period, there will be only one final disposition code.
Appendix E

Components and descriptions of each category of response rate calculation (for a sampling frame of housing units) (American Association for Public Opinion Research, 2016)

- To standardize the response rate calculations across countries, every country should group each sample element’s final disposition code into one of the following mutually exclusive and exhaustive categories:
  A. Interviews
  B. Non-interviews—Eligible
  C. Non-interviews—Unknown eligibility
  D. Non-interviews—Ineligible

A. Interviews

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete interviews</td>
<td>• Respondent has finished the interview.</td>
</tr>
<tr>
<td>Partial interviews</td>
<td>• The survey organization (in consultation with the coordinating center) may decide prior to the start of data collection to consider an interview to be a partial interview if at least some percent (e.g., 80%) of applicable or crucial/essential questions have been answered.</td>
</tr>
<tr>
<td>TOTAL INTERVIEWS</td>
<td>• Sum of interviews.</td>
</tr>
</tbody>
</table>

B. Non-interviews—Eligible

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Refusals</td>
<td>• It has been determined that there is an eligible respondent in the housing unit but either he/she or someone else refuses the interview request.</td>
</tr>
<tr>
<td>Non-contacts</td>
<td>• It has been determined that there is an eligible respondent in the housing unit but the interviewer cannot gain access to the building, no one is reached at the housing unit, or the respondent is never available when the interviewer attempts an interview.</td>
</tr>
<tr>
<td>Other</td>
<td>• It has been determined that there is an eligible respondent in the household (eligibility determined as of a particular date, e.g., the date that the household listing is taken) but at some time after the determination of eligibility, the respondent is unable</td>
</tr>
</tbody>
</table>
to complete the interview due to reasons other than a refusal or is unable to be reached after repeated attempts. For example, the respondent may have died, been incarcerated or hospitalized, or left the country.

- It has been determined that there is an eligible respondent in the household, but he/she does not speak any of the study language(s) or is permanently incapable of participating in the interview due to a physical or mental condition (e.g., senility, blindness, or deafness). Note: Sample elements may be considered ineligible if the target population is defined such that respondents who do not speak the study language(s) or respondents who are unable to hear are explicitly excluded from the target population to which the study plans to makes inferences.

- Any other eligible non-interview status.

<table>
<thead>
<tr>
<th>TOTAL NON-INTERVIEWS—ELIGIBLE</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sum of eligible non-interviews.</td>
<td></td>
</tr>
<tr>
<td>If the survey organization is unable to provide separate counts of each component but the survey organization can provide the total number of eligible non-interviews, use the total.</td>
<td></td>
</tr>
</tbody>
</table>

C. Non-interviews—Unknown eligibility

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
</tr>
</thead>
</table>
| Unknown if household/occupied housing unit | - The sample elements have not been attempted or worked (e.g., no interviewer is available in area or replicates are introduced too late to work all sample elements).
- Interviewer is unable to reach the housing unit due to weather or concerns about safety in a dangerous neighborhood.
- Interviewer is unable to locate the housing unit (e.g., inaccurate or inadequate address/locating information). |

| Unknown if eligible respondent is in unit/no screener completed | It has been determined that there is an eligible housing unit but the interviewer is unable to determine whether there is an eligible respondent in the unit. For example, a household member may refuse to complete the screener or no one is available to complete the screener when the interviewer visits the household. Note: |
These sample elements are not considered refusals, since only elements where it has been determined that there is an eligible respondent can be classified as refusals.

<table>
<thead>
<tr>
<th>Other</th>
<th>Any other status for which eligibility is unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOTAL NON-INTERVIEWS—UNKNOWN ELIGIBILITY</td>
<td>Sum of non-interviews of unknown eligibility</td>
</tr>
<tr>
<td></td>
<td>If the survey organization is unable to provide separate counts of each component, but the survey organization can provide the total number of non-interviews of unknown eligibility, use the total.</td>
</tr>
</tbody>
</table>

### D. Non-interviews—Ineligible

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not an eligible housing unit</td>
<td>The sample elements are out-of-sample housing units or housing units that are incorrectly listed in the address frame (e.g., housing units are outside the primary sampling unit in which they are thought to be located).</td>
</tr>
<tr>
<td></td>
<td>The sample elements are non-residential units (e.g., businesses, government offices, institutions, or group quarters).</td>
</tr>
<tr>
<td></td>
<td>Housing units are vacant on the date that eligibility is determined. Note: Sample elements may be considered eligible non-interviews if someone is present at the housing unit on the date that eligibility is determined, even if when the interviewer returns the household has moved and the unit is vacant.</td>
</tr>
<tr>
<td></td>
<td>Households are temporary, seasonal, or vacation residences (i.e., not the usual place of residence).</td>
</tr>
<tr>
<td>No eligible respondent</td>
<td>It has been determined that there is an eligible housing unit, but there is no eligible respondent in the unit. For example:</td>
</tr>
<tr>
<td></td>
<td>Residence with no one 18 years of age or older.</td>
</tr>
<tr>
<td></td>
<td>Respondent does not speak any of the study language(s) and the target population is explicitly defined such that respondents who do not speak the study language(s) are not considered part of the target population to which the study plans to make inferences (may also hold for physical or mental conditions, if the target population is explicitly defined to exclude persons who are blind, deaf, senile, etc.).</td>
</tr>
<tr>
<td></td>
<td>Respondent died before eligibility is determined.</td>
</tr>
<tr>
<td><strong>Other</strong></td>
<td><strong>TOTAL NON-INTERVIEWS—INELIGIBLE</strong></td>
</tr>
<tr>
<td>----------------</td>
<td>-----------------------------------</td>
</tr>
</tbody>
</table>
| • Respondent is incarcerated or hospitalized (i.e., institutionalized) at the time that eligibility is determined, and remains institutionalized throughout the data collection period. | • Sum of ineligible non-interviews.  
• If the survey organization is unable to provide separate counts of each component but the survey organization can provide the total number of ineligible non-interviews, use the total. |
| • Respondent is in a group/cell for which the quota has already been filled.  
• Any other ineligible non-interview status. |
Appendix F

Recording counts of response rate categories template (for a sampling frame of housing units) (American Association for Public Opinion Research, 2016)

- Use the template below to help determine the number (or weighted count, if appropriate) of sample elements finalized in each of the categories and, thus, the total number/weighted count of sample elements fielded. The total number of sample elements is the sum of all categories of the final disposition codes.

- First, enter the number of sample elements finalized as each given category component. If no sample elements are finalized as a particular category component, enter “0” in the “Count” column.

- Next, total the components for each category by entering the sum on the longer of the “Count” column lines.

- Finally, total the sums of each category by entering the overall sum on the last “Count” column line.

- Use the “Additional Information” column to provide any information that will assist in interpreting the figures provided, particularly the study’s definition of partial interviews or descriptions of “Other” classifications specific to the study.

<table>
<thead>
<tr>
<th>Category (with Components)</th>
<th>Count</th>
<th>Additional Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Interviews</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complete interviews</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partial interviews</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL INTERVIEWS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. Non-interviews—Eligible</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Refusals</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-contacts</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL NON-INTERVIEWS—ELIGIBLE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. Non-interviews—Unknown eligibility</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown if household/occupied housing unit</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown if eligible respondent in unit/no screener completed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL NON-INTERVIEWS—UNKNOWN ELIGIBILITY</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| D. Non-interviews—Ineligibility| | }
<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not an eligible housing unit</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No eligible respondent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TOTAL NON-INTERVIEWS—</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>INELIGIBILITY</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TOTAL NUMBER OF SAMPLE ELEMENTS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix G

Recording counts of response rate categories for additional eligible respondents template (for a sampling frame of housing units) ([American Association for Public Opinion Research, 2016](#))

- Use the template below to help determine the number (or weighted count, if appropriate) of additional respondents in each of the categories and, thus, the total number/weighted count of additional respondents. The total number of additional respondents is the sum of only the eligible categories of the final disposition codes; if a household was not eligible, no respondents—let alone additional respondents—were selected.

  - First, enter the number of additional respondents finalized as each given category component. If no additional respondents are finalized in a particular category component, enter “0” in the “Count” column.
  
  - Next, total the components for each category by entering the sum on the longer of the “Count” column lines.
  
  - Finally, total the sums of each category by entering the overall sum on the last “Count” column line.
  
  - Use “Additional Information” column to provide any information that will assist in interpreting the figures provided, particularly the study’s definition of partial interviews or descriptions of the “Other” classification specific to the study.

<table>
<thead>
<tr>
<th>Category (with Components)</th>
<th>Count</th>
<th>Additional Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Interviews</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complete interviews</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partial interviews</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL INTERVIEWS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. Non-interviews—Eligible</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Refusals</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-contacts</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL NON-INTERVIEWS—ELIGIBLE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL NUMBER OF ADDITIONAL RESPONDENTS</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Data Collection: Face-to-Face Surveys

*Julie de Jong, 2016*

**Introduction**

Many cross-cultural projects attempt to keep the mode of administration constant by choosing face-to-face data collection, where the survey questionnaire is administered, at least in part, by a survey interviewer. Generally, the face-to-face mode has the best sample coverage properties, highest response rates (and therefore possibly lower nonresponse bias), and does not require respondents to be literate. For a discussion of the advantages and disadvantages of the face-to-face mode of interviewing, see *Study Design and Organizational Structure*.

Before the advent of personal computing, face-to-face surveys were administered using a paper-and-pencil instrument (PAPI). However, laptops and other electronic instruments (e.g., tablets, smartphones, etc.) are now widely used in lieu of PAPI.

In order to collect comparable data, multinational, multiregional, and multicultural surveys ("3MC" surveys) must establish a standard data collection protocol. At the same time, the protocol will sometimes need to allow for modifications required by local norms, conditions, or customs.

The implementation of face-to-face surveys presents a number of logistical challenges not faced in other modes. This chapter first addresses issues pertaining to the face-to-face mode, regardless of the instrument used to collect the data (i.e., paper and pencil questionnaire, computerized instrument, etc.), and then presents considerations particular to each type of instrument.

**Guidelines**

**Goal**: To achieve an optimal 3MC data collection design by maximizing the amount of information obtained per monetary unit spent within the allotted time, while meeting the specified level of precision and producing comparable results, within the context of a face-to-face survey.

1. **Consider the following steps when conducting survey interviews using a face-to-face mode.** Surveys conducted by interviewers face-to-face share a number of common procedural steps.

   **Rationale**

   There are a number of important considerations when interviewers are contacting respondents in a face-to-face survey, whether the instrument is paper-based (PAPI) or computer-based (CAPI).
**Procedural Steps**

1.1 Contact local authorities for clearance for the interviewers to collect data at the sample site(s); if necessary, negotiate with local authorities or, in some cases, military authorities to gain access to sample areas.

1.2 Allow adequate time for interviewer recruitment and training
   1.2.1 Match interviewer and respondent characteristics (e.g., race, ethnicity, or gender) when cultural norms so dictate, and/or if there is reason to think that interviewer effects may occur depending on the social conditions (see Interviewer Recruitment, Selection, and Training).
   1.2.2 While interviewers might be relatively easy to recruit in some countries, in other places such as the Gulf States and Middle East region, researchers might face some challenges in recruiting qualified field interviewers (Gengler, 2013).

1.3 Take measures to ensure interviewer safety.
   1.3.1 Inquire about potential safety problems, such as civil unrest and high crime areas.
   1.3.2 Decide whether interviewers should travel in groups and be accompanied by security personnel.
   1.3.3 Have interviewers visit their work areas during the daytime before the first day of data collection. They should check for potential hazards and safe havens during this visit.
   1.3.4 Have interviewers tell their supervisors and family members when they plan to leave for the field, the location of the area, and when to expect them back.

1.4 Have interviewers carry the following items in the field to establish their legitimacy:
   1.4.1 Official identification from the survey organization.
   1.4.2 Official letters to local authorities describing the study, if appropriate.
   1.4.3 Other letters of permission or support from local authorities if appropriate and/or necessary given the local social context and governmental regulations.

1.5 Provide adequate transportation and accommodation for staff and supplies.
   1.5.1 If maps are unavailable or unreliable, consider the use of local guides or GPS instruments.
   1.5.2 Arrange to secure fuel and oil and to maintain the vehicles used by the field staff; this may present logistical problems in some countries where there are breakdowns in infrastructure.
1.5.3 Arrange for emergency transportation in the event that a field team member becomes ill or injured and needs immediate medical attention or it becomes unsafe to stay in an area.
1.5.4 Arrange for backup transportation.
1.5.5 Secure housing accommodations in more remote areas prior to fieldwork or have the team carry their own (e.g., tents or mobile homes).

1.6 If physical measurements are taken as part of the survey, check the cultural acceptance of taking such measurements.

1.7 Provide all members of the field staff with access to a reliable line of communication with their supervisor. This will allow them to report progress and problems.
1.7.1 Majority countries may have weak communication capacities, especially in rural areas.
1.7.2 Cellular or satellite phones may be a worthwhile investment for teams in remote areas.

1.8 Aim to conduct the interview in a setting which affords visual, physical, and auditory privacy.
1.8.1 Privacy is critical for keeping respondents' answers to the survey questions confidential.
1.8.2 Although complete privacy is ideal, it is impossible to achieve in some cultures. Interviewers should attempt to keep the interview as private as possible, while still respecting cultural norms. This may involve self-administration on more sensitive questions. See Guideline 4 below regarding self-administration in the context of a face-to-face interview. An alternative may be to keep any others present occupied while the targeted respondent completes the survey.
1.8.3 In some countries, it may be unacceptable to have an interviewer come to the respondent's home, or it may be unacceptable for an interviewer of the opposite sex to interview or enter the home of the selected respondent or informant. As noted above, this may necessitate interviewer-respondent gender matching.
1.8.4 Privacy increases the likelihood that respondents will answer honestly about sensitive behaviors, such as sexual practices or drug use, or about sensitive attitudes such as religion in some contexts. What is considered sensitive may vary among countries or cultures; administration practices may need to differ accordingly.
1.9 In order to reduce non-response in the face-to-face mode of interviewing, train the interviewers to make observations of the housing unit to assess likely at-home patterns.
1.9.1 Note that in some countries interviewers are not allowed to ask neighbors about targeted but not yet contacted respondents.

Lessons learned

1.1 Because responses to some survey questions can be affected by other individuals present during data collection, it is optimal—but not always possible—to conduct face-to-face surveys in private. In a face-to-face fertility survey of women in what is now Bangladesh, privacy was difficult to establish; most interviews took place in the presence of the respondent's mother- or sister-in-law. This may have affected responses to sensitive questions (Choldin, Kahn, & Ara, 1983).

1.2 Similarly, men in some parts of Africa were found to object to confidential interviews of their wives or children. The interviewers were instructed to conduct interviews in a place that was visible to the male heads of household but out of earshot (Chikwanha, 2005).

1.3 In some rural places it might not always be feasible to conduct an interview inside a home, and may have to take place outside and in a more public setting.

1.4 In other rural places, the survey interview is still a novel concept, making interview privacy difficult to attain. In the Chitwan Valley Family Surveys in Nepal, a survey interview often draws family members and even interested neighbors, who sit with the respondent and interviewer to listen in.

1.5 Analyses using data from nine countries participating in the World Mental Health Survey Initiative provided evidence that the presence of a third party during the survey interview process affected the reporting of sensitive information, but the effect is moderated by differences in social conformity and the cultural setting from country to country (Mneimneh, Tourangeau, Pennell, Heeringa, & Elliott, 2015).
2. **Consider the following steps when using a paper and pencil instrument (PAPI).**

*Rationale*

While the use of computerized technology has increased in survey administration, the paper-and-pencil instrument continues to be used by survey projects that lack the infrastructural capacity to adopt and maintain the necessary technology.

*Procedural Steps*

2.1 The paper instrument should be designed so that it is visually easy for the interviewer to administer. See *Instrument Technical Design* for further detail.

2.2 Develop a sample management protocol for use in the field by data collection supervisors. The protocol should include instructions for passing sampling units from one interviewer to another if the need arises, as well as the corresponding documentation of such transfers (see *Data Collection: General Considerations* 3.3 and sample management system).

2.2.1 Use a coversheet to track each sample element during the study (see Appendix A for an example of a paper coversheet).

2.2.2 Interviewers using paper coversheets have found that they work most efficiently if they sort the coversheets by (1) appointment time and (2) geographical location.

2.2.3 Consider efficient methods that allow interviewers to fill in coversheets and do household contacting at the same time. Filling in coversheet forms after making the contact has shown to be error prone.

2.3 Train interviewers to complete household enumeration and randomly select eligible members within the household unit (*Kulka & Weeks, 1988*) (see Appendix B for household enumeration and Appendix C for an example of a Kish table).

2.4 Develop a distribution procedure for supplies to interviewers in the field, including a surplus of paper questionnaires to ensure a continual supply.

2.5 Develop a protocol for transferring completed paper questionnaires from interviewers to field supervisors, and from field supervisors to the head office or other location where data entry will occur.
2.6 Develop a protocol for maintaining completed questionnaires and coversheets in a secure location to ensure protection of respondent confidentiality.

Lessons Learned

2.1 Paper questionnaires and other survey materials can be misplaced, stolen, or otherwise lost in the field. Document any such circumstances and develop a protocol to determine whether affected respondents will be recontacted for a repeat interview.

2.2 Plan for adequate storage, security, filing system to get back to interviews efficiently.

2.3 In certain countries, like Ghana, weather conditions such as high humidity can destroy paper questionnaires in storage.

2.4 Researchers administering a PAPI survey of business and social entrepreneurship in the Kingdom of Tonga with complex skip patterns used a detailed skip pattern map to simplify training and questionnaire preparation. They also developed a notation system on the actual questionnaire page to assist the interviewer (Frederick, 2012).

2.5 If there are multiple components to the questionnaire, consider using paper of different colors for each component (e.g., the coversheet in yellow, interviewer-completed survey in green, self-administered section in orange, etc.).

2.6 Alternately, if the questionnaire will be administered in several different dialects or languages within a country, consider printing each dialect/language on differently-colored paper.

2.7 Consider using heavy card stock or lamination for Show Cards and other paper-based instruments that will be used for multiple respondents.

2.8 If using an event history calendar or other unusually-sized instrument, allow for adequate printing time, particularly in countries where printing of odd-sized documents may be challenging. Researchers in Nepal report having difficulties in locating printing businesses with the capacity to print the large life-history calendars designed for administration.

2.9 Researchers administering a PAPI survey in the Kingdom of Tonga faced limitations in printing in the country itself, including the lack of
paper, printing cartridges, and water-resistant paper that could withstand moisture and travel. Researchers emailed the questionnaire to a specialty printer in New Zealand, and the printed questionnaires were sent via airmail back to Tonga for use (Frederick, 2012).

2.10 The cost of paper can be very expensive in some countries. If the survey instrument contains many skip patterns, there can be a lot of waste as well. For example, the PAPI version of the World Mental Health Composite International Diagnostic Interview (CIDI) 3.0 was about 400 pages in length, but contained numerous sections which began with a screener question and resulted in the respondent skipping the entire section(s) for which the questions were not applicable.

3. **If an electronic instrument will be used instead of a paper-based instrument, consider the following procedural steps.**

**Rationale**

As technology becomes more accessible and affordable, with use increasing worldwide, **computer-assisted personal interviewing** (CAPI) is a popular mode choice and is frequently used in lieu of PAPI. Laptop computers have generally been the instrument of choice for CAPI, but tablets, smartphones, and other handheld device are increasing in popularity.

**Procedural Steps**

3.1 If CAPI use is new to the study site, develop an introduction strategy for both local collaborators and study respondents.

3.1.1 Involve local collaborators in study design if possible to facilitate its adoption. The clinical and administrative staff in a rural Kenyan health center aided in the identification of appropriate data and formatting of the paper and electronic data recording interfaces. This helped reduce fears and distrust of computers and engaged the clinical staff in the clinical research project (Diero et al., 2006).

3.1.2 In settings with limited technology, computerization can stimulate survey respondent interest and add legitimacy to the interviewers. Interviewers might also be more motivated to use technology in such setting. However, at the same time and in certain cultures, the use of technology can raise suspicion among respondents (Paudel, Ahmed, Pradhan, & Dangol, 2013).
3.2 Assess technical experience at the data collection firm.

3.2.1 Critical staff should have adequate language competency. Programs interested in incorporating technology into their activities need to hire bilingual staff as trainers and programmers to improve understanding of how to use the chosen technology, and to facilitate design and analysis activities, as most technology specifications are available only in universally-used languages such as English.

3.2.2 The data collection organization needs to have technical expertise to create the questionnaire, provide technical support for interviewers, manage the in-flow and out-flow of data, manage databases, and run quality control checks. While some of these tasks could be outsourced, building local capacity is always recommended for continuity and long-term goals.

3.3 Assess available infrastructure in the study country.

3.3.1 If the data collection organization requires data to be transmitted on a regular basis for quality control, and reliable Internet connectivity needs to be in place, evaluate WiFi and other Internet connection capabilities across the geographic areas covered by the sample. Even though a country’s major cities may have adequate Internet capabilities allowing for regular data transfer, rural areas may present more challenges.

3.3.2 Interviewers and field office staff need to have access to reliable electrical power sources for the interviewing as well as communication devices (e.g. mobile phones). Interviewers might need to carry multiple batteries for their devices if they are visiting areas with limited power supply. Interviewers could also be instructed to use other methods for charging batteries including in-car chargers such as cigarette-lighter adapter or portable generators (Shirima et al., 2007; Byass et al., 2008)

3.3.3 If the need to revise the questionnaire during data collection arises, computerization and connectivity allows for an easy transmission of updated questionnaires to interviewers or respondents without the need for reprinting, mailing, or personal pick-up of material. Moreover, avoiding printing any material at, before, or during production is environmentally friendly.

3.4 Choose and procure the necessary primary and auxiliary equipment.

3.4.1 Primary equipment

- There must be a good fit between the project and the technological tool. Handheld devices may be more appropriate for smaller or simpler questionnaires, and,
because of their size, devices such as smartphones are not as suitable for collecting open-ended responses (Escandon, Searing, Goldberg, Duran, & Monterrey Arce, 2008).

- Purchasing equipment and accessories locally can facilitate more efficient servicing of equipment than if equipment is purchased internationally (Paudel et al., 2013).

- Although new technology may be more expensive if purchased locally within less-developed countries, the cost saved in shipping, delays, and in-country technical support can more than compensate for that difference.

- If equipment is not available locally; however, most hardware is available through collaborators in industrialized countries or can be ordered directly via the Internet.

- Consider ordering an excess supply of batteries and extra equipment (e.g., several extra laptops) in case of equipment malfunction.

3.4.2 Auxiliary equipment

- Decide on a backup and uploading process (SD cards, flash drives, automatic uploading to central system, etc.).

- Data synchronization between a mobile device and a central computer can be very time consuming in a rural, remote setting. In a survey in Zanzibar, a mobile device was used to collect data, store and copy the data from the SD card to the central computer (Thriemer et al., 2012).

- The back-up system must be carefully developed to handle possible transitions or losses. In a root-cause analysis from a survey using PDAs in Bolivia, poor back-up protocol, due to programmer error, precluded researchers from interpreting the data (Escandon et al., 2008).

- If possible, at least two separate central back-up systems should be developed, in addition to having back-up on the unit itself (i.e., memory cards) and a communal archiving system.

3.5 Select appropriate data collection software.

3.5.1 Additional attention should be given to non-Latin languages (i.e., Chinese, Arabic, Russian, etc.) when selecting technology and programming software. Not all software packages can support non-Latin script.

3.6 Select an appropriate electronic sample management system.

3.6.1 If an electronic sample management system is used, coordinating centers may play a role in monitoring
fieldwork. See Study Design and Organizational Structure for details.

3.6.2 The electronic sample management system should permit interviewers to be able to sort the sample respondents by (1) appointment time and (2) geographical location.

3.7 Develop and test the CAPI instrument
3.7.1 Allow for sufficient time and budget for computerized specifications in the preproduction phase (House & Nicholls, 1988)
3.7.2 Consider using paper documents for certain aspects of the survey. For example, interviewers in China using handheld computers reported that it was overly time-consuming to read the full consent form on a small screen (Wan et al., 2013).

3.8 Develop a distribution system for supplies to the field.
3.8.1 Develop procedures for storage and transport of equipment.
3.8.2 Interviewers who are traveling long distances, through difficult terrains, or weather conditions find it easier to carry their laptop or even smaller devices (tablets) to conduct their interviews than carrying cumbersome paper questionnaires (Paudel et al., 2013).

3.9 Develop procedures for use and maintenance of technology in the field.
3.9.1 Charge batteries daily to mitigate data loss due to battery discharge. Instruct interviewers to verify daily that batteries are charged.
3.9.2 Provide interviewers with a reliable electrical source to charge both CAPI instrument and mobile phones batteries so that interviewers can contact supervisors in the event of equipment malfunction. Communication is necessary for possible instrument troubleshooting and monitoring team progress. Most technical issues are simple user errors that can be resolved with a short discussion with the supervisor.
3.9.3 Backup plans need to be designed and implemented in case of power outages, especially in resource-constrained environments.
3.9.4 It is possible for data to be lost because of hardware or software malfunction and for equipment to be lost or stolen during fieldwork. Researchers need to establish protocols for preventing and handling such situations.
3.9.5 Decide whether interviewers should be provided with paper copies of the questionnaire or some material to take notes in-case of equipment failure (Onono, Carraher, Cohen, Bukusi, & Turan, 2011). Some studies choose not to provide paper
versions because they do not want to encourage use of alternate paper instrument by interviewers.

3.9.6 It is crucial to have local informatics experts for development and custom integration of databases, continued support, and adaptation to new applications. Specifically, a programmer with experience in database and systems design, implementation, and maintenance is recommended, and this resident expertise is available in most if not all countries (Avilés, Ortega, Kuan, Coloma, & Harris, 2007).

3.9.7 Equip interviewers with accessories that are needed for protecting and maintaining the equipment such as laptop bags, screen covers, sleeves, rain shields, etc.

3.9.8 Ask interviewers sign term of use agreement detailing equipment’s ownership and responsibilities.

3.10 Management of data files during the field period.

3.10.1 The electronic data audit trail provides important paradata and should be uploaded and backed up as well. Determine what will happen to paradata in case of equipment failure during interview.

3.10.2 Lack of electricity and/or Internet connection can lead to delays in the backup and uploading process. For example, a survey in Kenya experienced delays in immediate transfer of data collected due to electrical instability, and data often could not be backed up in the field so was only backed up once a week at the study office (Onono et al., 2011).

3.11 Develop strategies to increase privacy

3.11.1 Though interviewer-administered computerization can in general increase the level of respondent privacy, the novelty of it in some cultures might attract bystanders, and interviewers may need additional training on how to request and achieve privacy in such situations (Paudel et al., 2013). DHS interviewers in Nepal found they often had to make extra effort to maintain privacy, which usually demanded more time to administer the questionnaire (Paudel et al., 2013). See also Ali et al. (2010).

3.11.2 Reading computer screens under direct sunlight can lead to difficulty in administering an interview and limit the options for confidential interview space. This can be a particular concern when asking sensitive questions related to sexual behavior and domestic violence.

3.12 Devote adequate time to interviewer training for CAPI-specific issues. When using CAPI, interviewer training is a two-step process, requiring technical training focused specifically on the survey
Instruct interviewers on how to introduce technology to the survey population, especially in settings where exposure to technology is more limited. This could be done by collaborating with community leaders who could act as liaisons announcing the survey and the use of technology to their community members.

3.12.2 Instruct interviewers on how to explain the use of technology to respondents during the consent process (e.g., that recording will or will not be disabled).

3.12.3 Provide training on how to handle, label, care, transport, and store equipment properly. This is especially important in contexts where technology is more novel.

3.12.4 Instruct interviewers on steps to take in case of equipment failure and theft.

3.12.5 Instruct interviewers on password use, stylus if needed, how to access the questionnaire, enter responses, and insert and remove any memory cards used.

3.12.6 Operational instructions should be in study site language and not only in English (Wan et al., 2013).

3.12.7 If paper questionnaires will be available in the event of equipment malfunction, training on the PAPI instrument is essential as well.

3.12.8 When using technology, there can be a tendency for interviewers to focus on the technology rather than the respondent, which should be addressed during interviewer training.

3.12.9 Allow interviewers ample time to practice administering the questionnaire to increase comfort with the flow of questions. Interviewers are more likely to lose track of where they are in the sequences of questions because they can see only one screen at a time, and familiarity with the instrument can decrease difficulty (Groves & Mathiowetz, 1984; House & Nicholls, 1988; Couper, 2000).

3.12.10 If using an electronic sample management system, train interviewers to complete household enumeration and randomly select eligible members within the household unit.

3.12.11 Although interviewers must be trained in the use of the specific computer program, it is crucial to devote adequate time to training on other important interpersonal aspects of survey implementation (Groves et al., 2009).
3.13 Develop interviewer management procedures for use in the field.
   3.13.1 Interviewers must have fast and regular communication of field staff with team leaders and technical support staff. This is necessary for troubleshooting and monitoring team progress.
   3.13.2 Information technologies allow implementation of a system of work ownership if all personnel are assigned a code for database entry, supervision, and analysis to maintain logs controlling data management and information flow.

Lessons Learned

3.1 Technology can be adopted even in resource-poor countries, leading to improvements in efficiency and data collection capabilities.
   3.1.1 Researchers successfully conducted a Demographic and Health Survey (DHS) in Nepal using tablet PCs. The connection to the central network took, on average, one minute, and data transfer to the server in Kathmandu took approximately 5 to 7 minutes. In contrast, in the past, paper-based surveys had to be sent to Kathmandu via pouch mail or hand-carried, which took days or even weeks. The use of CAPI reduced data collection time by one month compared to the previous survey completed by PAPI (from 6.5 months to 5.5 months). However, there were some security concerns with carrying these tablets and storing them especially in remote areas because some interviewers had to stay in community members' homes. Enforcing joint responsibility for theft of, or damage to, the tablet PCs among the interviewer teams helped to ensure security of the tablets during transport and storage. For example, interviewers were trained to lock and be aware of their tablet PCs at all times, even during meal and rest times (Paudel et al., 2013). And, in a Peruvian survey, handheld computers were inserted into a wooden and Styrofoam clipboard to shield them from possible damage and to conceal them (Bernabe-Ortiz et al., 2008).
   3.1.2 In cross-cultural surveys such as the World Mental Health (WMH) Initiative, some participating countries have been unable to implement technology-based survey instruments due to infrastructural constraints. However, the WMH Coordination Centre made the decision that those countries which can, should use technology (in this case, CAPI), as the advantages outweigh the methodological concerns of non-comparability. Other experimental studies have found few significant differences in survey estimates (Baker, Bradburn, & Johnson, 1995; Couper, 2000). The WMH Organization’s current recommendation is to challenge where CAPI can and cannot be used. For example, in 2003, Columbia was able to
implement the WMH survey with great success using CAPI. Countries that used PAPI in the most recent WMH surveys expressed the wish that they had more strongly pursued CAPI, especially because of quality control and complexity of survey instrument (Pennell et al., 2008).

3.1.3 Researchers should be aware that mode differences can occur in unanticipated ways. In a meta-analysis of studies from the United States, Canada, the United Kingdom, and Italy comparing data from PDAs to PAPI, the results favor handheld computers over paper and pencil for data collection among study participants, but the data are not uniform for the different outcomes. Handheld computers appear superior in timeliness of receipt and data handling (four of four studies) and are preferred by most subjects (three of four studies). On the other hand, only one of the trials adequately compared adherence to instructions for recording and submission of data (handheld computers were superior), and comparisons of accuracy were inconsistent between five studies (Lane, Heddle, Arnold, & Walker, 2006).

3.1.4 The availability of information and communication technologies for direct data transfer has the potential to improve the conduct of research, and, especially, public health research, in resource-poor settings. Because of shortened data entry time in a vaccination survey in Zanzibar through use of CAPI, transition time to vaccination and subsequently to disease surveillance was shortened (Ali et al., 2010). As technology continues to evolve, research on its impact on survey data collection should continue.

3.1.5 In a study by Thriemer et al. (2012), a PDA-based survey in Tanzania resulted in an estimated 25% reduction in cost, compared to a paper-based survey. Elimination of questionnaire printing costs is even more significant if multiple languages/versions are needed in a country because multiple versions can be programmed into the platform (Onono et al., 2011). In another effort to reduce costs, researchers found that sending an excess supply of batteries to study sites helped decrease use of PAPI and its associated additional costs (Onono et al., 2011).

3.1.6 The use of technology can greatly increase the efficiency through which data from multiple data collection modes can be linked. Current Smartphone capabilities allow for scanning barcodes on respondent records, which has the potential to further effectively link data from multiple sources, such as completed surveys, signed letters of consent, medical charts, biomarker records, etc. (Aviles et al., 2007; Thriemer et al., 2012).
3.1.7 The use of CAPI platforms can extend potential working hours. Because CAPI can be used in low-light situations, interviewers can work during evening hours, otherwise a challenge with paper questionnaires in settings with frequent power outages.

3.2 In non-western settings, interviewers have generally reported a preference for CAPI instruments. Examples include the following from across the world:

3.2.1 CAPI was successfully implemented in a survey of malaria morbidity in Gambia, where interviewers reported a preference for CAPI over PAPI in terms of amount of work, number of errors, length of interviews, and ease of transport (Forster & Snow, 1992).

3.2.2 Handheld computers were used for a tobacco use survey in a hard-to-reach population in China where most interviewers stated a preference for handheld computers for future surveys (Wan et al., 2013).

3.2.3 In a survey in Zanzibar, acceptability of PDA use was high among staff not familiar with computers or PDAs and after an initial training period, none of the users was interested in returning to paper-based data entry (Thriemer et al., 2012).

3.2.4 In a survey in Bolivia, interviewers reported that using PDA to administer interviews stimulated their own interest in working on the survey (Escandon et al., 2008).

3.3 Use of technology has been well-received by respondents.

3.3.1 In a survey in Tanzania using PDAs, most respondents who expressed their opinions about the use of PDAs had something positive to say. For example, a 30-year old man with primary-level education said, "I was very happy to see a computer as it was my first time to see it. It simplified recording of our responses." An elderly man expressed his appreciation of having learnt what day of the week he had been born (Shirima et al., 2007).

3.3.2 In a Demographic and Health Survey in Nepal, respondents were curious about being interviewed using the tablet PCs. The interviewers perceived a high level of respect and enthusiasm from respondents, and they felt that respondents viewed them as technical employees with higher education. This was an unanticipated, but encouraging finding, especially because of respondents' limited exposure to computers. However, at the same time and in certain cultures, the use of technology can raise suspicion among respondents and although acceptability of the tablet PCs was high, there were a few cases of skepticism. As part of the informed consent
process, respondents were informed that the interview would not be video- or audio-recorded and that the recording feature had been disabled on the tablet PCs. However, a few respondents were still concerned (Paudel et al., 2013).

3.3.3 Analyses from a survey in rural south Kenya using PDAs found a reduction in refusals, attributed to the perception of respondents that the PDA was more secure (Onono et al., 2011).

3.3.4 Due to the increasing use of mobile phones and other similar technologies in day-to-day life, operating a computerized questionnaire on a handheld device might be more familiar to respondents with little or no experience in the use of computers (Ali et al., 2010).

3.4 Allow for adequate project preparation before beginning fieldwork. 
3.4.1 Do not underestimate the additional time needed for preparation for both initial adoption and continued use of technology. In a survey in Burkina Faso, researchers reported underestimating the amount of work required to program questionnaires, and as a result failed to maximize the use of some of the available options for input checking and other real-time quality control procedures. Village names, for example, were implemented as a text-entry field, but would have been better as a drop-down list to avoid ambiguities of spelling, etc. Combinations of input checks, plus quality control measures at the stage where data were downloaded to portable computers in the field, should have picked up concerns at an earlier and remediable stage (Byass et al., 2008).

3.4.2 Having local trained personnel is essential. Using a “train the trainers” model, technical and supervisory staffs in a public health survey in China were able to develop the questionnaire and complete the programming with minimal assistance from technical experts from the coordinating center. When problems occurred, the Chinese technical experts could then provide immediate technical guidance and trouble-shooting to interviewers and other staff (Wan et al., 2013).

3.4.3 It can be difficult to repair equipment in country. Aviles et al., (2007) recommend the implementation of preventative maintenance program.

3.4.4 Consent letters mentioning the use of technology can be helpful in reducing non-response. In a survey in Tanzania using PDAs, most respondents said that they had noted the PDA after its mention in the consent letter. Several interviewees appreciated the interviewer having introduced
them to the technology during the consent procedure (Shirima et al., 2007).

3.5 When using CAPI, particularly with interviewers previously unfamiliar with computerized instruments, consider the following with regards to interviewer recruitment, training, and management.

3.5.1 Experience suggests that interviewers with little education and no experience in the use of a computer are easily able to use handheld devices for survey administration (Ali et al., 2010). With increasing use of mobile phones and other similar technologies, operating handheld devices, downloading data, and recharging batteries are becoming increasingly familiar concepts.

3.5.2 Although use is increasing, however, plan for adequate time for interviewer training. In a survey in Bolivia using PDAs, interviewers wanted additional practice time because of previously limited experience with the technology, and particularly more instruction on the use of a stylus as keyboards on handheld devices can be cumbersome (Escandon et al., 2008).

3.5.3 Analyses of inter-observer accuracy and performance revealed a considerable range in a survey in Burkina Faso. Some interviewers clearly worked faster with the PDAs than others, though these were not necessarily those who covered the greatest number of households per day worked. However, those who carried out interviews relatively quickly were generally also those who made the least input errors. In surveys of this kind, where competence in local languages is an important factor, there are often not many options in terms of who can be recruited as interviewers (Byass et al., 2008).

3.5.4 Training on proper handling and care of the equipment is also very important, particularly in a rural context where the equipment has to be transported through rough terrain, the power supply is not stable, and unexpected rain is a concern. In the DHS survey in Nepal, teams were provided with generators, rain shields, umbrellas, and several other items to manage these challenges. Enforcing joint responsibility for theft of, or damage to, the tablet PCs among the interviewer teams helped to ensure security of the tablets during transport and storage. With proper care and maintenance, tablet PCs (and portable generators) can be reused in future surveys, resulting in additional cost savings over the long term (Paudel et al., 2013).
3.6 The use of CAPI is not without its technologically-related challenges. Project staff should be aware of the possibility for corrupted date/time stamps because of equipment malfunction. In a survey using PDAs in Kenya, researchers found that if the PDA lost power, it automatically reset the clock, which had effects on pregnancy data that was collected. Particular caution should be used if data is time-sensitive as in this case (Onono et al., 2011).

3.6.1 Project staff should be aware of the possibility for corrupted date/time stamps because of equipment malfunction. In a survey using PDAs in Kenya, researchers found that if the PDA lost power, it automatically reset the clock, which had effects on pregnancy data that was collected. Particular caution should be used if data is time-sensitive as in this case (Onono et al., 2011).

3.6.2 If using CAPI, the concurrent use of paper files for portions of the survey can lead to logistical challenges. Diero et al. (2006) used PDAs to follow patients who visited a rural Kenyan health center. The data entry program did not allow for entry of text field notes by the research assistants, who had to use a paper notebook for such notes. This can cause a disconnect between these text notes and the patient data to which they referred.

3.7 Researchers have used several methods for maintaining respondent confidentiality and ensuring data security when using CAPI.

3.7.1 Data can be copied and automatically saved to a SD card, after which interviewers are unable to retrieve or change an entry, with no record of the entry was retained on the PDA (Onono et al., 2011; Shirima et al., 2007). In case of equipment loss, it is then impossible to access and see the data on the SD card without a password and the requisite software. In a survey in Kenya, when one PDA was stolen in political violence, two interviews were lost on the SD card, but respondent confidentiality was maintained because of security protocols in place (Onono et al., 2011).

3.7.2 In a survey in Tanzania, data were downloaded to the laptop computers and daily summary reports produced to evaluate the completeness of data collection. Data were backed up at three levels: (i) at the end of every module, data were backed up onto storage cards in the PDA; (ii) at the end of every day, data were downloaded to laptop computers; and (iii) a compact disc (CD) was made of each team’s data each day (Shirima et al., 2007).

3.8 As smartphones become more ubiquitous in daily life, their use in survey research is expected to increase. Findings from a recent survey examining the usability of smartphones versus tablets in Kenya generally favor tablets over smartphones (Hughes & Haddaway, 2014). Highlights from the study include:

3.8.1 Confidence and comfort in typing dependent on past experience with device and touchscreens
3.8.2 Interviewers felt more likely to accidentally select options on phones
3.8.3 Interviewers admitted to not scrolling completely through questions/responses on phones
3.8.4 Interviewers felt more professional with tablets
3.8.5 Interviewers felt safer with phones because of the smaller size compared to the tablets, which attracted unwanted attention
3.8.6 Smartphones were associated with more typing errors
3.8.7 Long open-ended questions and long numeric strings are difficult

4. If the questionnaire includes items of a sensitive nature, consider administering these questions in a self-administered module during the face-to-face interview.

Rationale

Evidence suggests that increasing privacy during an interview can improve the accuracy of reporting such topics in surveys (Turner et al., 1998; Turner et al., 2002; Krawczyk et al., 2003), but achieving privacy in non-western settings varies considerably between countries (Mneimneh, 2012). For a face-to-face interview, consider administering the sensitive sections in a self-administered questionnaire (SAQ). Research indicates that respondents in an interviewer-administered, non-private setting tend to misreport information perceived to be sensitive. For example, respondents might underreport undesirable or private information such as drug use or illegal status; and they might over-report desirable information such as voting.

Many surveys include potentially sensitive questions about both respondent behavior and attitudes concerning such topics as sexual behavior and contraceptive use, substance abuse, violence, and politics. And, in non-western settings, these delicate topics are particularly susceptible to social desirability bias. However, asking sensitive questions in an SA format has the potential to decrease bias and achieve more accurate reporting.

Procedural Steps

4.1 Assess the literacy of the target population and choose the most appropriate instrument for the SAQ.
4.1.1 The SAQ can be a paper questionnaire given to the respondent to self-complete. The paper-based SAQ should not have complex skip patterns, and the target population should have adequate literacy levels.
4.1.2 The SAQ can take the form of computer-assisted self-interviewing (CASI), where respondents use a technology platform (i.e., a laptop, tablet, smartphone, etc.) and complete the entire questionnaire, or a specific section of the questionnaire, independently. The technology therefore facilitates the administration of a complex instrument, much like CAPI facilitates administration for the interviewer.

4.1.3 Audio-CASI (A-CASI) has the advantages of CASI, but can be particularly helpful in low-literacy settings. In A-CASI, respondents listen to an audio track recording of each survey question using a headset and move through the survey at their own pace. If illiterate, survey respondents can be instructed to push color-coded buttons on a touch screen or mini-keyboard, or have graphical representations of answer categories to indicate their response to each question (see Instrument Technical Design, Appendix F for an example).

4.1.4 If using A-CASI, assess whether the setting would benefit from gender-matching in terms of the audio voice used. That is, if the recording presented to female respondents should be a female voice, while male respondents are presented with a recorded male voice.

4.2 When designing an SAQ instrument, consider the following:

4.2.1 Be mindful of survey length. Longer surveys administered using an SAQ mode may have more missing data both because of lack of interviewer probing and lack of the pressure respondents feel to cooperate with the interviewer (Hewett, Erulkar, & Mensch, 2004a).

4.2.2 Develop interviewer instructions for explaining the SAQ to the respondent.

- The detail of instructions will differ by mode, with CASI and A-CASI necessitating more explanation than a paper-based SAQ, particularly in low-literacy settings.
- If an SAQ is utilized for reasons of increased respondent confidentiality, then this rationale should be explained to respondents.
- Develop a protocol for interviewer behavior during the interview, particularly concerning the extent to which interviewers should be encouraged to help the respondent or otherwise interact with the respondent. All interactions should be documented.
- Consider adding questions at the end of the interview to assess respondents’ perceived ease of use, privacy, and truthfulness.
4.2.3 When using CASI and A-CASI modes, attention to details that facilitate the respondent experience can lead to increased data quality.

- Consider disabling the screen saver and monitor power-saver settings on the device so that screens do not go blank if a participant takes additional time to answer a question (NIMH, 2007).
- Graphical and/or audio representations of the response process can help guide the respondent through the interview. In a survey in India using A-CASI, the entry of a response was marked by the change in the color of the corresponding response bar on the screen to grey, along with a “beep” sound. A “Thank you” screen indicated the end of the survey (Bhatnagar, Brown, Saravanamurthy, Kumar, & Detels, 2013).
- If a participant did not answer a question after approximately 60 seconds, consider repeating the question and/or programming additional text. The additional text can be programmed to appear encouraging participants to answer the item(s) in a truthful manner (NIMH, 2007).
- If used, the keyboard should be user-friendly. Keyboard options can be limited to responses (e.g. YES, NO, and numbers) and larger color-coded keyboard keys could be used. Additional keyboard shortcuts to replay questions can also be marked.
- Text on the computer screen should be large enough to be easily legible for respondents.
- In an A-CASI survey in India, neither the question nor the response texts were displayed on the screen to ensure privacy and confidentiality for the respondents (Bhatnagar et al., 2013).
- Touchscreens on A-CASI instruments can be particularly helpful for less-educated populations (Lara, Strickler, Olavarrieta, & Ellertson, 2004).

4.3 Additional technologies for SAQ mode in a face-to-face interview continue to emerge, including video-computer-administered self-interview (V-CASI) (Kissinger et al., 1999; Krysan & Couper, 2003). If planning to use an SAQ, investigate the most recent literature available for further guidance.

Lessons Learned

4.1 The use of novel technology, particularly in non-Western settings, can motivate respondents to participate.
4.1.1 In a comparison of paper SAQ vs. self-administered PDA questionnaires on sexual behavior, South African adolescents reported more favorable attitudes toward the PDA mode (Jaspan et al., 2007; Seebregts et al., 2009).

4.1.2 End-of-questionnaire items measured high respondent-perceived truthfulness in a South African survey about sexual behavior (Beauclair et al., 2013) and greater preference for A-CASI compared to other modes, primarily because of perceived increased confidentiality and privacy, as well as the novelty of technology (van de Wijgert, Padian, Shiboski, & Turner, 2000; Bhatnagar et al., 2013; Lara et al., 2004; Hewett, Mensch, & Erulkar, 2004b; Gutiérrez & Torres-Pereda, 2009).

4.1.3 However, if technology is unfamiliar to the population, it may cause concern about the project activities. In a study using A-CASI in rural Kenya, interviewers and supervisors reported that the presence of computers heightened the animosity and opposition of the community to the project activities. Rumors spread that the survey was the work of devil worshipers and that interviewers were collecting the names of adolescents who would later be abducted. Many respondents believed that the computers collected information for the government. Also, respondents were angry that expensive equipment was brought into resource-starved community during a time of drought. Misinformation spread throughout the region before interviewers even entered some sampling units. Some residents thought that the computer was having a “conversation” with the respondents, despite insistence that the computer voice was taped. And, in the initial A-CASI protocol respondents’ answers were read back to them after each question for verification, a protocol which needed to be discontinued because some respondents perceived the computer to be “talking to them”, resulting in decreased perceptions of confidentiality (Hewett et al., 2004a).

4.2 Use of an SAQ mode can impact the length of time needed for interviewer administration depending on setting and demographics. The HIV/STD Prevention Study found that surveys using A-CASI generally took longer to administer than CAPI in China, Peru, India, and Zimbabwe. However, A-CASI took less time in Russia, where the participants had more exposure to technology and were of a younger age (NIMH, 2007).

4.3 There is evidence that using A-CASI is feasible in non-western settings.
4.3.1 The NIMH Collaborative HIV/STD Prevention Trial Group conducted a feasibility study comparing results from surveys using CAPI and A-CASI in China, India, Peru, Russia, and Zimbabwe (NIMH, 2007). Despite the varying levels of literacy and exposure to computers by country, most study participants reported that it was easy to enter their answers into the computer, that they felt comfortable doing so, and that they preferred the computer rather than an interviewer for answering questions about topics such as sexual behavior and drug and alcohol use, or had no preference. Most participants gave the same responses on both their A-CASI and CAPI interviews.

4.3.2 While A-CASI has generally been feasible in non-western settings, however, ease of use can vary by socio-demographic characteristics.
- Older and unemployed respondents report increased difficulty with A-CASI (Beauclair et al., 2013), as do less educated respondents (van de Wijgert et al., 2000).
- Women with little education (primary school or less) had considerably more problems using the computer keyboard, reading the computer screen, and correcting mistakes than women in higher educational groups (also Gutierrez & Torres-Pereda, 2009).

4.4 In regions where there are multiple languages and dialects, use of A-CASI can facilitate the interview process. A completely self-administered questionnaire can ease the logistical challenges in the field of matching a respondent with an interviewer who has the necessary language capabilities.

4.5 Use of A-CASI can lead to improvements in data quality.
4.5.1 A-CASI is a more standardized method of assessment than CAPI. Using CAPI, interviewers may use probes beyond the standard set even though they are instructed not to do so.
4.5.2 Unlike a paper-based SAQ, use of A-CASI leads to fewer data entry errors and missing data because the skip patterns are programmed into the computer and are executed as the interview is administered (van Griensven et al., 2006; Langhaug et al., 2011).
4.5.3 In a comparison study on topics related to HIV/AIDS in three cities in Vietnam, respondents assigned to A-CASI had lower item refusal rates than those assigned to a face-to-face interview or a paper-based-SAQ (Le & Vu, 2012).
4.5.4 There is evidence that using A-CASI has the potential to improve data quality through the reduction of missing data. Studies of mode differences in South Africa and Thailand have
found that those respondents assigned A-CASI had less missing data than those assigned to a paper-based SAQ (Jaspan et al., 2007; van Griensven et al., 2006).

4.5.5 A short respondent-training session prior to the administration of A-CASI can improve data quality. A survey of young women in Malawi utilized headphones and an external color-coded minikeypad, with a red key to replay the question, a green key to go on to the next question, and a yellow key to skip a question. For dichotomous questions, respondents were instructed to press 1 for yes and 2 for no. Prior to the A-CASI main survey, each respondent completed three “practice” questions to evaluate her understanding of the interview process, for example, “Are you a male or a female?” For each practice question, the correct answers were previously entered by the interviewer to serve as a check against the respondent’s entry. Respondents were not able to proceed to the main interview until they were able to answer all three practice questions correctly (Mensch, Hewett, Gregory, & Helleringer, 2008).

4.6 On the other hand, use of A-CASI can bring challenges to data quality as a result of decreased interviewer interaction.

4.6.1 Respondents may not understand skip patterns or other aspects of the survey but are reluctant to ask the interviewer for direct assistance given the hands-off nature of A-CASI (Lara et al., 2004).

4.6.2 An SAQ on sensitive topics may also lead to reluctance to engage the interviewer in a related question about completing the survey because, in using A-CASI, there can be an underlying perception that the topic is too delicate to discuss outright.

4.6.3 A study using SAQ in Tanzania found that about 7% of respondents selected only the first or the last response categories in a section for which such a response pattern would be inconsistent. This bias was associated with females, those less educated, and those more geographically remote (Plummer et al., 2004a).

4.6.4 Mode can impact data quality because of inconsistency in editing. A comparative survey of young women in Malawi found more consistent reporting in the face-to-face mode than in the A-CASI mode. Researchers speculated that, contrary to protocol, the face-to-face interviewers may have edited the questions for consistency post-hoc, whereas such editing was not possible in the A-CASI mode by respondents.
4.7 A-CASI is often used with the a-priori expectation that privacy for the respondent will result in increased reporting of more sensitive behaviors, with the related implication that this reporting is indeed accurate (Groves et al., 2009; Couper, 2005). However, meta-analyses using data from non-western settings are inconclusive on whether SAQ modes increase accuracy of sensitive behaviors.

4.8 Results from a meta-analysis of face-to-face and A-CASI modes in studies on sexually transmitted infections and associated behaviors in Brazil, Vietnam, Thailand, Kenya, India, Russia, Zimbabwe, Malawi, China, and Tanzania demonstrate that overall, A-CASI methods are not consistently associated with a significant increase in reporting of sensitive behaviors, but trends can be seen in certain contexts. In general, increased reporting in A-CASI has been associated with region (Asia), setting (urban), and education (secondary education) (Phillips, Gomez, Boily, & Garnett, 2010).

4.9 In contrast, a meta-analysis by Langhaug, Sherr, and Cowan. (2010) of 26 studies in developing countries on sexual behavior demonstrated that, in general, A-CASI can significantly reduce reporting bias. The results of this review as well as findings from other researchers (cited below) show that the relationship and success of novel interviewing methods has proved complex in a low- and middle-income country context and researchers should be aware of the mode differences that can result, depending on the study topic and social context (NIMH, 2007; Jaspan et al., 2007; Rathod, Minnis, Subbiah, & Krishnan, 2011; Langhaug et al., 2011; Lara et al., 2004; Mensch, Hewett, & Erulkar, 2003; Mensch et al., 2008; Mensch et al., 2011; Hewett et al., 2004b; Plummer et al., 2004a; Plummer et al., 2004b; Potdar & Koenig, 2005; Minnis et al., 2009; Jaya, Hindin, & Ahmed, 2008).
Appendix A

Cover-sheet (example from the University of Michigan’s Institute for Social Research)

Final Result Code: __________
Date of Final Result (dd/mm/yyyy): ___ / ___ / ______
Length of interview: _________ Total calls: ______
Length of edit:
Appendix B

Household enumeration table (example from the University of Michigan’s Institute for Social Research)

<table>
<thead>
<tr>
<th>HOUSEHOLD ENUMERATION</th>
<th>RESPONDENT SELECTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>11 a. Household Member’s First Name</td>
<td>11 b. HH Member’s Relationship to Informant</td>
</tr>
<tr>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>A</td>
<td>M</td>
</tr>
<tr>
<td>L</td>
<td>M</td>
</tr>
<tr>
<td>E</td>
<td>M</td>
</tr>
<tr>
<td>S</td>
<td>M</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>E</td>
<td>F</td>
</tr>
<tr>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td>A</td>
<td>F</td>
</tr>
<tr>
<td>L</td>
<td>F</td>
</tr>
<tr>
<td>E</td>
<td>F</td>
</tr>
</tbody>
</table>

Instructions for household enumeration table

Column 11a (Household Member’s First Name): List all members of the household, beginning with the informant. Note that males are listed in the upper portion of the table and females in the lower portion.

Column 11b (Household Member’s Relationship to Informant): Record each household member’s relationship to the informant (e.g., husband or wife, son or daughter, mother or father, brother or sister, friend, etc.).

Column 11d (Age): Record each household member’s age.

Column 11e (Language Spoken): This column may or may not be included, depending upon the study requirements.

Column 11f (Eligible): Place a check mark in this column if, based upon the information in columns 11a-11e, the household member meets the eligibility criteria for the study.

Column 11g (Person Number): Assign a sequential number to each eligible household member. Begin by numbering eligible males from oldest to youngest, continue by numbering eligible females from oldest to youngest.
Column 11h (Selected R): Count the number of eligible persons in the household. Find that number in the Kish table in the "If the Number of Eligible Persons is:" column. The selected respondent will be the household member with the "Person Number" corresponding to the "Interview the Person Numbered:" column in the Kish table.
Appendix C

Random within house selection techniques

**Kish (1949) selection tables**

<table>
<thead>
<tr>
<th>SELECTION TABLE A</th>
<th>SELECTION TABLE B₁</th>
<th>SELECTION TABLE B₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>If the number of eligible persons is:</td>
<td>Interview the person numbered:</td>
<td>If the number of eligible persons is:</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>6 or more</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SELECTION TABLE C</th>
<th>SELECTION TABLE D</th>
<th>SELECTION TABLE E₁</th>
</tr>
</thead>
<tbody>
<tr>
<td>If the number of eligible persons is:</td>
<td>Interview the person numbered:</td>
<td>If the number of eligible persons is:</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>6 or more</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SELECTION TABLE E₂</th>
<th>SELECTION TABLE F</th>
</tr>
</thead>
<tbody>
<tr>
<td>If the number of eligible persons is:</td>
<td>Interview the person numbered:</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>6 or more</td>
<td>5</td>
</tr>
</tbody>
</table>

The tables provide unbiased estimates by giving each respondent a weight based on the number of adults in the household. This guarantees that the selection within a household is random for a combined total random sample across the housing units (addresses) that were randomly selected in the first place.
Instructions for Kish tables

1. Assigning Kish tables to the sample file: one of the twelve tables is randomly selected and assigned to the first line in the sample file. The series of twelve is then run through twice, assigning tables to the sample lines. Then again a table is randomly selected and the series is run through twice. This procedure is repeated until all sample lines have an assigned Kish table.

2. Household listing: a household listing of eligible adults (age 18 and over) who reside in that household is taken at each of the sample addresses. Usually the males are listed first in order of decreasing age, and then the females in the same order.

3. Using Kish tables: the table assigns a number to each member of the household listing. Sample Kish tables are shown above. In the first column the interviewer would circle the total number of eligible persons. The corresponding number in the second column of the Selection Table denotes the person selected to be interviewed.
Data Collection: Telephone Surveys

Julie de Jong, 2016

Introduction

If researchers wish to have survey interviews carried out by an interviewer, but face-to-face interviews are not possible, conducting interviews via telephone either through a landline or mobile telephone can be an alternative. Multinational, multiregional, and multicultural survey (“3MC” surveys) use different standards to determine whether telephone penetration is adequate in a study country. For example, the Gallup World Poll generally uses a telephone survey only in countries where telephone coverage represents at least 80% of the population (Gallup, 2015). Telephone interviews are generally less costly than face-to-face methods, and can be completed in a shorter amount of time. However, response rates are generally lower and depending on the available sampling frame for a country, a rigorous telephone-administered sample design can be difficult to develop. See Sample Design for a discussion of the challenges and limitations of a telephone-based frame and sample design.

As discussed in Data Collection: General Considerations, 3MC surveys sometimes employ mixed modes, depending on individual country constraints. However, it is important to note that mode effects may occur if the survey is carried out by telephone in some countries and face-to-face in others (see Study Design and Organizational Structure for discussion on mode effects).

Virtually all questionnaires administered by interviewers in telephone surveys are completed using an electronic computer-based instrument to record survey responses. This data collection mode is most commonly referred to as computer-assisted telephone interviewing (CATI). These guidelines assume that the interviewer will be using a computer-based instrument and will refer to the mode as CATI.

For additional discussion on the advantages and disadvantages of telephone surveys, see Study Design and Organizational Structure.

Guidelines

Goal: To achieve an optimal cross-cultural data collection design by maximizing the amount of information obtained per monetary unit spent within the allotted time, while meeting the specified level of precision and producing comparable results, within the context of a telephone survey.

1. Develop the computer-based system(s) that the interviewers will use to administer telephone interviews.
**Rationale**

Interviewers can conduct telephone interviews from either a central location or remotely. Software systems can be used to distribute sampled telephone numbers, to dial telephone numbers, to manage call records, and to record survey data. When using CATI, it is crucial to design and implement a system that interviewers can use to reliably collect survey data.

**Procedural steps**

1. **Decide whether interviewers will work in a centralized and/or decentralized location.**
   1.1 Many survey research firms conducting telephone interviews maintain a “telephone lab,” which is a central calling center where center supervisors oversee a variable number of interviewers. Each interviewer has access to the electronic instrument and records responses directly in the electronic file. Interviews can be monitored in real time.
   1.1.2 Sometimes interviewers work from other locations while having access to the electronic system set up by the survey research firm.

2. **Develop a system and protocol for sample release management, including how cases will be transferred between interviewers when necessary.**

3. **Develop a protocol for dialing sampled telephone numbers.** Some projects may use CATI systems that can dial telephone numbers automatically, while other projects may elect to have interviewers dial telephone numbers manually. In some countries it is against the law to use automation to dial specific types of telephone numbers (e.g., in the United States, it is illegal to use automation to dial mobile numbers). If using automation, be familiar with the local laws about its use.

4. **Consider the cost structure for telephone calls in each study country.** In the United States, respondents are responsible for the cost of incoming telephone calls on mobile telephones. However, in the Persian Gulf, for example, there is no charge and interviewers based in Nepal were able to telephone Nepali migrant workers living in Gulf countries for a migration survey without any cost to the respondents (Ghimir, Williams, Thornton, Young-DeMarco, & Bhandari, 2013).

5. **Decide which telephone number and name will be displayed to the respondents in the caller ID, and whether the telephone number**
should be available if people call back the number.

1.6 Develop an electronic survey instrument used to record survey responses. There are numerous CATI software packages. However, it is also possible to use a web-based survey instrument, which may not be as suitable for more complex projects but is less expensive. Electronic survey instruments in a telephone survey share many of the same requirements as electronic survey instruments administered in the face-to-face mode. For in-depth discussion of these elements, see Guideline 3 in Data Collection: Face-To-Face Surveys.

Lessons Learned

1.1 While survey mode can affect survey responses, studies are not unanimous in the direction of the effect observed.

1.1.1 A survey of HPV awareness and knowledge, including sexual behavior, was conducted in Singapore, with half participating via CATI and half through an interviewer-administered face-to-face interview. Few differences between survey modes were found in the information disclosed (Smith et al., 2009).

1.1.2 A study in India evaluating accuracy of health data collection through several different interfaces found that telephone interviewing had greatest accuracy in phone interviews when compared to electronic forms on PDAs and text messaging (Patnaik, Brunskill, & Thies, 2009).

1.2 CATI can be particularly useful in a panel study setting, especially when there is frequent contact with respondents. Experiences vary by country, however.

1.2.1 In a study of farmers in Tanzania, researchers gave respondents pre-paid mobile phones for the duration of the field period so that they could receive a phone call from an interviewer and complete a survey every three weeks over a ten-month period, resulting in a high quality dataset (Dillon, 2012).

1.2.2 Researchers distributed mobile phones to female sex workers in India for use in a diary study on sexual behavior, which resulted in high response rates and high-quality data (Bradley et al., 2012).

1.2.3 Researchers on a panel study in South Sudan using CATI found that response rates were affected by irregular fluctuations in the mobile network (Demombynes, Gubbins, & Romeo, 2013).
1.3 Beyond the traditional CATI mode, interviewing via text message has been recently used. In this mode, the interviewer sends individual survey questions by text to the respondent, who sends his or her responses back by text to the interviewer (West, Ghimire, & Axinn, 2015; Lau, Lombaard, Baker, Eyerman, & Thalij, 2016).

2. **Train interviewers on interviewing strategies specific to telephone interviewing.**

*Rationale*

The nature of the interaction between the interviewer and the respondent depends on the mode of data collection. Some interviewing strategies that are accessible in a face-to-face mode, such as interpretation of body language, are not possible to implement over the telephone, contributing in part to lower response rates and potential for non-response bias. However, there are certain telephone-specific strategies that researchers can introduce to assist interviewers in completing telephone interviews.

*Procedural steps*

2.1 Consider the social context of the study country when hiring interviewers to administer a telephone survey, and whether selection of interviewer based on gender or other characteristics will affect response rates. See Lessons Learned 2.1 below as well as Interviewer Recruitment, Selection, and Training for additional discussion of interviewer recruitment considerations.

2.2 Develop an introduction appropriate for the interviewer to read upon contact with the respondent.

2.2.1 The introduction is especially important and may differ depending on cultural norms, and the way the opening unfolds between the interviewer and respondent may have significant implications for both survey non-response and data quality (Couper & Groves, 2002). The context of the interview can dictate identification procedures and pace of interview.

2.2.2 Establishing and maintaining rapport is especially important in achieving a telephone survey. Particular care should be taken in the translation stage to ensure an interviewer script that does not violate cultural norms involving politeness and linguistic encoding of status and social distance (Kleiner & Pan, 2006).

2.2.3 The introduction can be particularly critical in achieving cooperation in some countries. Previous respondent exposure to the telephone as a survey mode can differ across countries,
and there can be discomfort in sharing personal information over the phone (Hughes, 2004).

2.2.4 In countries where there are linguistic differences depending on actors’ social status, translations must also recognize that interviewers and respondents are strangers and cannot rely on visual cues to establish social distance and appropriate linguistic level, necessitating the opportunity for some social interaction at the beginning of the survey to establish such social distance.

Lessons Learned

2.1 Gender norms of the study country can have a significant impact on response rates in CATI surveys.

2.1.1 In France, researchers have found that female interviewers generally have higher refusal rates in telephone surveys (Verger, Baruffol, & Rotily, 2001).

2.1.2 In Nepal, a highly gendered society, women generally prefer to speak to other women, and men to men, even over the telephone. However, in a CATI survey using Nepali-based interviewers contacting (mostly male) Nepali migrant workers in Persian Gulf countries, researchers obtained high response rates using predominantly female interviewers, because of the cultural perception that women would not call a male unless it was an important matter (Ghimire et al., 2013).

2.1.3 There is also anecdotal evidence that male respondents in the highly gendered countries in the Middle East are more likely to participate in a telephone survey when contacted by a female interviewer.

2.2 Immediate identification by name is standard telephone practice in the United States, but is uncommon in China (Kleiner & Pan, 2006).

2.3 Acceptable pace of the interview introduction can vary across even otherwise similar cultural contexts. For example, an examination of reaction to phone calls in Hong Kong and Beijing found that Beijing residents were more resistant to a fast-paced, business-like telephone conversation when compared to those from Hong Kong (Pan, Scollon, & Scollon, 2002). Similarly, a comparison of Greeks and Germans showed that Greeks prefer social interaction before reaching the main point of a telephone conversation, while Germans prefer to discuss the main point immediately (Pavlidou, 1994).

2.4 Acquiescence bias differs across cultures and can be particularly problematic in a telephone survey where otherwise difficult issues can be exaggerated. For example, in many Asian cultures, people
tend to avoid “no” answers to yes/no questions, particularly when there is an asymmetrical relationship between speakers as in a survey interview (Kleiner & Pan, 2006).

2.5 Introductory scripts can differ dramatically across cultures. For example, in Chinese, the use of expressions like “please” and “thank you” are not normally used in daily conversation and imply a large social distance between speakers. The mandated repetitive use of such words in a survey among Chinese speakers would be detrimental, particularly in a telephone survey where rapport is especially important, in sharp contrast to a survey in American English, where such phrases are acceptable and expected (Pan et al., 2002).

3. Decide whether a subset of survey questions would best be collected in a self-administered section of the interview.

**Rationale**

Interviewer-administered telephone interviewing is subject to social desirability biases similar to those in face-to-face interviewing. Interactive Voice Response (IVR) is a telephone mode where the computer plays recordings of the questions over the telephone to respondents who then respond by using the keypad of the telephone or saying their answers aloud. IVR can be used as a self-administered mode (SAQ) to administer a portion of an interview, otherwise conducted by CATI, which is particularly sensitive in nature and where accuracy might improve without the presence of an interviewer. It can also be used exclusively as a self-administered mode (SAQ), with the computer automatically telephoning the respondent and then completing the questionnaire (see Data Collection: Self-Administered Surveys) for further discussion of IVR in a completely self-administered mode.

**Procedural Steps**

3.1 Design the IVR system so that it is technically well-integrated into the CATI system in use by the project and that switching from the CATI to the IVR system is straightforward for the interviewer.

3.2 Decide whether to program the IVR system as touchtone, voice input, or a combination of the two.
3.2.1 When deciding on the programming, consider the target population. Studies in rural India and Botswana found that respondents with less education and lower literacy do better with touchtone, and cited privacy for touchtone preference as well (Kuun, 2010; Patel et al., 2009).
3.2.2 A study in Pakistan found that a well-designed speech interface was more effective than a touch-tone system for respondents regardless of literacy level (Sherwani et al., 2009).

3.3 Devote sufficient time to the development of a high-quality IVR system to maintain respondent interest and continued cooperation.

3.3.1 The IVR system must have a high quality recording, as the respondent is likely to break off the survey if quality is poor.

3.3.2 See Oberle (2008) for a guide to the development of an IVR system and the associated speech characteristics which need consideration.

Lessons Learned

3.1 Consider the voice used for recording.

3.1.1 In a health helpline project in Botswana, researchers employed a well-known local actress for the IVR recording, and users reacted very positively (Kuun, 2010).

3.1.2 Depending on the social context, using an IVR recording of a male for male respondents and of a female for female respondents may elicit more accurate reporting, particularly of sensitive information.

3.2 Plauche, Nallasamay, Pal, Wooters, and Ramachandran (2006) developed an innovative approach to the challenge that dialectical variation and multilingualism poses to speech-driven interfaces for IVR in India, applicable to other settings as well. In their approach, people from specific villages are recorded during interactions, and their speech is semi-automatically integrated into the acoustic models for that village, thus generating the linguistic resources needed for automatic recognition of their speech.

3.3 A survey of teachers in Uganda resulted in a number of useful considerations when designing an IVR system to improve response rates and data quality (Lerer, Ward, & Amarasinghe., 2010).

3.3.1 The IVR call began with the immediate information that “This is a recorded call from Project X. You are not talking to a real person.”

3.3.2 The IVR call provided very specific instructions about whether to use keypad or to speak.

3.3.3 Respondents were initially confused by the automation of the IVR system. Researchers had better results when using a chime to get respondents’ attention before the automated voice gave instructions.
3.3.4 Leveraging conversational and turn-taking conventions of normal conversation in the IVR system lead to more success than detailed instructions in eliciting desired user behavior.

3.3.5 An IVR system which projected a loud voice, with prompts recorded as if the speaker were using a poor cell connection, resulted in a survey that was easier for respondents to follow.

3.3.6 When producing the IVR recording, use slow speech to get slow speech – respondents will emulate the voice, and resulting data will be easier to understand.

3.3.7 The IVR recording included 3 seconds of silence before the recorded speakers says “thank you” and moves onto next question, which was reported as well-received by respondents.
Data Collection: Self-Administered Surveys

Julie de Jong, 2016

Introduction

Fully self-administered questionnaires (SAQ) are not as common as interviewer administered surveys in the context of multinational, multiregional, and multicultural surveys (“3MC” surveys). However, as surveys become more costly to administer using interviewers, whether face-to-face or by telephone, more researchers are considering SAQ modes. SAQ modes include mail surveys, web surveys, and Interactive Voice Response (IVR) surveys conducted via telephone.

An important element of the self-administered mode is that there is, by definition, no interviewer involved. As discussed in detail in Survey Quality, interviewer error can contribute significantly to total survey error. In removing the interviewer from the equation, survey quality can improve. This may be particularly true if the survey topic is sensitive. Self-administered modes can also be effective when the when privacy during the survey interview is difficult to obtain.

However, the absence of an interviewer also demands a carefully designed survey instrument that is easy for the respondent to complete. Because there is no interviewer present, there is no one to assist the respondent in understanding instructions or to provide encouragement to complete the questionnaire. Differences in literacy levels among countries should also be considered in the questionnaire design phase of self-administered instruments (see Questionnaire Design for further details).

In addition, because of the lack of interviewer-respondent interaction, non-response is more difficult to assess and it is a challenge to disentangle the effects of noncontact, refusal, and a poor sampling frame. For example, non-response to a mail survey may result from misdirected mail that never arrived at the sample respondent’s house, misplaced mail within the respondent’s house, initial willingness to complete the survey but subsequent forgetfulness, unwillingness to complete the questionnaire (i.e., a refusal), or any number of issues. And, in a multi-person household, it may be impossible to identify who the actual respondent was. Therefore, when designing an SAQ, it is crucial to implement strategies to maximize survey quality.

For further discussion on the advantages and disadvantages of self-administered surveys, see Study Design and Organizational Structure. For additional information on sample design and related challenges for self-administered modes, see Sample Design].
Guidelines

Goal: To achieve an optimal cross-cultural data collection design by maximizing the amount of information obtained per monetary unit spent within the allotted time, while meeting the specified level of precision and producing comparable results, within the context of a self-administered survey.

1. When a mail survey using a paper-based instrument will be sent to respondents, develop the questionnaire and protocols with consideration that the survey must be straightforward for respondents to self-administer.

Rationale

Concerns about response rates, length of surveys, and quality of data have all resulted in a reduction in the use of mail surveys in recent years. However, Dillman and others argue that high quality mail surveys, with close attention to detail, can result in accurate data (Dillman et al., 2007; Dillman, Smyth, & Christian, 2014). The mail survey is becoming more widespread as the cost of interviewer-administered surveys increases. If a mail survey is the chosen mode of data collection, consider the following steps when developing the instrument.

Procedural Steps

1.1 Assess the postal system in the study country and use it to develop a timeline for data collection that is realistic given the local context. In a 3MC survey, there are often differences in postal reliability, cost, possible carriers, and timeliness.

1.2 When designing materials (letters, questionnaires, etc.) that will be mailed to the respondent, assess the following:
   1.2.1 Literacy levels among the target population
   1.2.2 Use of languages and/or regional dialects other than the country’s official language(s) and any implications for the feasibility of a self-completed questionnaire. Indeed, there are some languages and dialects that do not have a written form.

1.3 Determine how data entry of returned mail questionnaires will occur. Data entry can occur manually but it is more efficient to use optical or intelligent character recognition software, wherein the computer will read and code response from paper questionnaires.

1.4 Before mailing out the paper questionnaire, consider sending a well-written advance letter to legitimize the survey and reassure and motivate potential respondents. Most effective is a carefully drafted,
simple, short letter (Couper & de Leeuw, 2003; Lynn, Turner, & Smith, 1997; Dillman, 2000).

1.5 Develop a cover letter to include with the paper questionnaire, introducing the research study and explaining the purpose of the survey, instructions on how to complete the instrument, and organization contact information for any questions the respondent might have.

1.6 Develop an instrument appropriate for the mode and target population, keeping in mind that there will be no interviewer present to assist with the survey administration.
   1.6.1 Assess the literacy of the target population and adjust the text for comprehension if necessary.
   1.6.2 Place instructions clearly next to the survey questions to which they correspond.
   1.6.3 Make the layout of the instrument visually appealing and question order easy to follow. Use visual elements (e.g., brightness, color, shape, position on page) in a consistent way to define the desired path through the questionnaire (Jenkins & Dillman, 1997; Groves, et al., 2009).
   1.6.4 Use skip patterns only when absolutely necessary. Include clear instructions for skip patterns and reinforce with visual and graphical cues, such as boldfacing and arrows.
   1.6.5 Limit the number of open-ended questions.
   1.6.6 Ask only one question at a time. Combining multiple items into one question places a heavy cognitive burden on respondents and can impact data quality.

1.7 Provide clear instructions for returning the completed survey to the research organization or other point of collection. Adequate postage should be provided on the envelope so as not to incur cost to the respondent.

1.8 Develop a sample management system to process completed paper questionnaires and develop procedures for its execution.

1.9 Institute protocols to protect respondent confidentiality. It is common for research organizations to assign a unique identification number to each sampled household's questionnaire for sample management purposes as questionnaire are mailed back to the office. This ensures that if a paper questionnaire is lost in the mail or is not otherwise returned to the survey organization, the respondent's answers cannot be linked to their identity by a third party.
1.10 Develop a protocol for addressing non-response, including how many attempts to reach respondents by mail and/or other possible methods will be made.

Lessons Learned

1.1 Because a mail survey is self-administered without an interviewer present, it is crucial that the layout and design of the questionnaire elements is clear and easy to follow, and that instructions are visibly marked. Often, the first page of a mail survey contains a lengthy set of instructions which respondents generally skip or do not retain when completing the questionnaire, argue Jenkins and Dillman (1997). They advise the placement of relevant instructions be directly where they need to be.

1.2 A recent mail survey in Siberia, which varied experimental factors across random subgroups of respondents, achieved greatest response rates when official university letterhead was used in correspondence, when there was an incentive offered, and when a larger number, versus a smaller number, of contacts with the respondent were attempted (Avdeyeva & Matland, 2012).

1.3 Expected response rates for mail surveys will differ by country. For a limited set of studies examining cross-national differences in response rates, see Couper and de Leeuw (2003), de Heer (1999), Hox et al. (2002), and Smith (2007).

2. When administering a survey via the web (i.e., the Internet), develop the questionnaire and protocols with consideration that the survey must be straightforward for respondents to self-administer.

Rationale

Internet penetration has been steadily increasing worldwide in recent years. Given the increased costs of interviewer-administered surveys, many researchers are turning to the use of web-based surveys to reach respondents when an adequate sample frame is available. Web surveys should be designed so respondents can easily access and complete the survey.

Procedural Steps

2.1 Assess each study country’s technological infrastructure to select software appropriate for use, depending on instruments prevalent in
the study country, for the development, distribution, and completion of the web survey.

2.1.1 Assess Internet speed and reliability in the study country and potential impact on ease of web survey use by respondents and design the survey to fit the country’s bandwidth limitations.

2.1.2 Determine which web browser(s) fully supports the web-based survey instrument and communicate this to the respondent. Consider including a link to download a specific web browser to facilitate the respondent’s participation in the web survey.

2.1.3 Consider that respondents will likely use different devices to access the survey, including desktop computers, laptop computers, tablets, smartphones, and other electronic devices. The web survey should be able to be completed on a web browser, regardless of the type electronic device. See Instrument Technical Design for additional information on preparing style sheets appropriate for multiple devices.

2.1.4 Plan for adequate programming and testing time on multiple devices. For example, software that may be compatible with Android devices may have glitches in iOS (Apple) devices.

2.2 Determine how respondents will be invited to participate in the web survey.

2.2.1 Before disseminating the link to the web-based survey instrument, consider sending a well-written advance letter to legitimize the survey and reassure and motivate potential respondents. Most effective is a carefully drafted, simple, short letter (Couper & de Leeuw, 2003; Lynn et al., 1997; Dillman, 2000).

2.2.2 Mode of invitation will be limited by the respondent contact information available from the sample frame. For example, a web survey using a sampling frame consisting solely of email addresses will not be able to send an invitation via postal mail because of the lack of a mailing address.

2.3 Determine how respondents will gain access to the survey. Dillman (2000) proposes providing a PIN number for limiting access only to people in the sample. Another option is to provide each respondent with a unique Internet link to the survey, which is linked to the respondent’s sample id.

2.4 Develop a concise introduction to be presented at the start of the web survey, introducing the research study and explaining the purpose of the survey, instructions on how to complete the survey, and organization contact information for any questions the respondent might have.
2.5 Develop and test the web survey, keeping in mind that there will be no interviewer present to assist with the survey administration.
   2.5.1 Assess the literacy of the target population and adjust the text for comprehension if necessary.
   2.5.2 The first question should be an item that is likely to be interesting to most respondents and easy to answer.
   2.5.3 Place instructions alongside the survey questions to which they correspond.
   2.5.4 Make the layout of the instrument visually appealing.
   2.5.5 Program any skip patterns used directly into the instrument, relieving the respondent from navigational decisions.
   2.5.6 Keep the survey as brief and engaging as possible. The longer the questionnaire and the greater the number of screens, the more likely the respondent will not finish the questionnaire (MacElroy, 2000).
   2.5.7 Limit the number of open-ended questions.
   2.5.8 Ask only one question at a time. Combining multiple items into one question places a heavy cognitive burden on respondents and can impact data quality.
   2.5.9 Make prompts, particularly those asking for the respondent to correct an answer, helpful, polite, and encouraging.
   2.5.10 Decide whether respondents can navigate backwards to revisit and/or revise previous survey items and responses.
   2.5.11 See Instrument Technical Design for additional guidance on the layout and technical design of the web survey.

2.6 Decide whether respondents will be permitted to complete the questionnaire in more than one session, allowing for the data to be saved in the interim, and program the instrument accordingly.

2.7 Institute protocols to protect respondent confidentiality.
   2.7.1 Ensure that electronic transmission of the data from the respondent’s computer to the survey firm collecting the data is secure.

2.8 Select an appropriate electronic sample management system and develop procedures for its execution. If an electronic sample management system is used, coordinating centers may play a role in monitoring fieldwork. See Study Design and Organizational Structure for details.

2.9 Determine which paradata will be collected. Paradata from web surveys can be used to enhance respondents’ experience or to understand more about the respondents and how they interact with the web survey (Couper, 2008). See Paradata and Other Auxiliary Data for more information and examples.
2.10 Develop a protocol for addressing non-response, including how many attempts to reach respondents by email and/or other possible methods will be made.

Lessons Learned

2.1 Web surveys are often used in subsequent waves of panel surveys following an interviewer-administered baseline study, and can be a practical and cost-effective mode choice. In such cases, the respondent is familiar with the study and strategies to minimize non-response can be executed via phone, mail, and even in-person visits because complete contact information is generally available.

2.2 With adequate design, web surveys can achieve response rates comparable to non-web surveys.

2.2.1 A randomized telephone/web mode experiment in a Swiss election study found that the use of an incentive in a web survey produced response rates comparable to those from the telephone survey which also included incentives. The web survey was much less costly, even accounting for the cost of incentives, than for the telephone survey (Lipps & Pekari, 2013).

2.2.2 However, like 3MC surveys conducted in other modes, web surveys can produce difference response rates across countries. A comparison of data collected through a web survey from Italy, France, Turkey, and the U.S. showed that France had the highest overall refusal rate, but low item nonresponse for those who did participate. Italy and the U.S. had response rates and low item non response. Respondent in Turkey had the lowest contact and response rates, and the highest item nonresponse for sensitive questions (Ackermann, Ecklund, Phillips, & Brulia, 2016).

2.3 Internet censorship occurs at the national level in at least several non-Western countries, such as China and Iran. If planning a survey in a country where censorship occurs, consider the survey topic and technical programming and determine whether the web is an acceptable form of data collection for the particular study country.

2.3.1 Censorship by certain governments can impact the types of questions that are permitted on a web survey questionnaire.

2.3.2 Censorship can impact response rates due to confidentiality and security concerns among respondents.

2.3.3 If the study country engages in censorship, consider the location of the server hosting the survey, and whether the study respondents will be able to access the server in its host country; that is, whether the server website IP address is
accessible from the study country.

2.4 Software and website vendors can restrict access by users in other countries. Regardless of any government censorship, verify that respondents in the study country can access the survey.

2.5 Smartphone apps are currently being used for time-use surveys. For example, a research study in the Netherlands is using a smartphone app to collect time use data in combination with auxiliary data. By requiring respondents to install an app, rather than access a website to complete the survey, researchers can guarantee that respondents will visually see the instrument exactly as the researchers intended. The app does not need permanent Internet access as completed survey data is stored and transmitted as Internet access permits (Sonck & Fernee, 2013).

3. When administering a survey using IVR, develop the questionnaire and protocols with consideration that the survey must be straightforward for respondents to self-administer.

Rationale

IVR can be an effective mode for administering a survey to a population where telephone accessibility is adequate and particularly when the survey topic is sensitive. However, as with mail and web surveys, the absence of an interviewer necessitates careful instrument design and field execution.

Procedural Steps

3.1 Determine which IVR software will be used to carry out the survey, including whether the IVR system will accept incoming phone calls from respondents to complete the survey and/or will initiate outgoing telephone calls to respondents to complete the survey.

3.2 Determine how respondents will be invited to participate in the IVR survey. Mode of invitation will be limited by the respondent contact information available from the sample frame.
   3.2.1 If postal addresses are available, respondents can receive an invitation with a telephone number to call to participate.
   3.2.2 If email addresses are available, respondents can receive an invitation and telephone via email.
   3.2.3 If only telephone numbers are available, the invitation to complete the IVR will occur by telephone.
3.3 If an automated dialing system will be used to initiate contact with the respondent, assess any legal restrictions in place that apply to the use of such systems in the study country.

3.4 Develop a concise introduction to be presented at the start of the IVR survey, introducing the research study and explaining the purpose of the survey, instructions on how to complete the survey, and organization contact information for any questions the respondent might have.

3.5 Decide whether to program the IVR system as touchtone, voice input, or a combination of the two.

3.5.1 When deciding on the programming, consider the target population. Studies in rural India and Botswana found that respondents with less education and lower literacy do better with touchtone, and cited privacy for touchtone preference as well (Kuun, 2010; Patel et al., 2009).

3.5.2 A study in Pakistan found that a well-designed speech interface was more effective than a touch-tone system for respondents regardless of literacy level (Sherwani et al., 2009).

3.6 Devote sufficient time to the development of a high-quality IVR system to maintain respondent interest and continued cooperation.

3.6.1 The IVR system must have a high quality recording, as the respondent is likely to break off the survey if quality is poor.

3.6.2 See Oberle (2008) for a guide to the development of an IVR system and the associated speech characteristics which need consideration.

3.7 Select an appropriate sample management system and develop procedures for its execution.

3.7.1 If an electronic sample management system is used, coordinating centers may play a role in monitoring fieldwork. See Study Design and Organizational Structure for details.

3.8 Develop a protocol for addressing non-response, including how many attempts to reach respondents by telephone and/or other possible methods will be made.

**Lessons Learned**

3.1 Consider the voice used for recording.
3.1.1 In a health helpline project in Botswana, researchers employed a well-known local actress for the IVR recording, and users reacted very positively (Kuun, 2010).

3.1.2 Depending on the social context, using an IVR recording of a male for male respondents and of a female for female respondents may elicit more accurate reporting, particularly of sensitive information.

3.2 Plauche, Nallasamy, Pal, Wooters, and Ramachandran (2006) developed an innovative approach to the challenge that dialectical variation and multilingualism poses to speech-driven interfaces for IVR in India, applicable to other settings as well. In their approach, people from specific villages are recorded during interactions, and their speech is semi-automatically integrated into the acoustic models for that village, thus generating the linguistic resources needed for automatic recognition of their speech.

3.3 Consider an alternate mode for first contact to inform respondent of impending IVR survey, such as SMS or other mailing. In a study in rural Uganda, the IVR survey call was preceded by an SMS message 24 hours prior, about the upcoming call. In a pretest, respondents who didn’t receive the text were unable to make sense of the later survey call (Lerer, Ward, & Amarasinghe, 2010).

3.4 A survey of teachers in Uganda resulted in a number of useful considerations when designing an IVR system to improve response rates and data quality (Lerer et al., 2010).

3.4.1 The IVR call began with the immediate information that “This is a recorded call from Project X. You are not talking to a real person.”

3.4.2 The IVR call provided very specific instructions about whether to use keypad or to speak

3.4.3 Respondents were initially confused by the automation of the IVR system. Researchers had better results when using a chime to get respondents’ attention before the automated voice gave instructions.

3.4.4 Leveraging conversational and turn-taking conventions of normal conversation in the IVR system lead to more success than detailed instructions in eliciting desired user behavior.

3.4.5 An IVR system which projected a loud voice, with prompts recorded like the speaker was using a poor cell connection, resulted in a survey that was easier for respondents to follow.

3.4.6 When producing the IVR recording, use slow speech to get slow speech – respondents will emulate the voice, and resulting data will be easier to understand.
3.4.7 The IVR recording included 3 seconds of silence before the recorded speakers says “thank you” and moves onto next question, which was reported as well-received by respondents.
Data Collection References


Hulme, D. (2007). *Integrating quantitative and qualitative research for country case studies of development*. The work was part of the program of the Economic and Social Research Council, Global Poverty Research Group. (No. GPRG-WPS-063).


Li, L. (2011) Retrieved on May 19, 2016, from [http://www.lukeexperiments.com/content/index/media/Luke_Li_Eye_Tracking_Methodology.pdf](http://www.lukeexperiments.com/content/index/media/Luke_Li_Eye_Tracking_Methodology.pdf)


Cross-Cultural Survey Guidelines

Do not distribute or reprint without permission


Further Reading


