CHAPTER 7

Telephone-Based Research Methods in Disaster Research

SANDRO GALEA, MICHAEL BUCUVALAS,
HEIDI RESNICK, JOHN BOYLE,
DAVID VLAHOV, and DEAN KILPATRICK

Telephone-based research methods have gained prominence as a potential tool in the armamentarium of the disaster researcher. Telephone-based methods allow for the rapid sampling and assessment of large populations and offer particular advantages for researchers interested in the consequences of disasters. This chapter provides a practical introduction to these methods. We focus here on how such research can be conducted and identify key issues that researchers should consider before and during the course of such research. We discuss the potential advantages and disadvantages of telephone-based research with reference to the situations when such methods may be more or less appropriate. We refer to "telephone-based research methods" as study methods that involve recruitment and interviewing of study participants through telephones. It is possible to recruit participants by telephone and interview them in person or using other methods, or vice versa, to recruit participants using other methods (e.g., off lists containing particular subgroups such as college students) and then interview these persons by telephone. We use the terms *telephone sampling* or *telephone interviewing*, respectively, to refer to these special instances.

DESIGN ISSUES IN TELEPHONE SURVEYS

Sampling

Choosing a Sampling Frame

The first practical question when considering the application of a telephone-based research method in the postdisaster context is the particular sampling frame that is the target of the survey. It is the goal of all research to obtain information on persons who are of interest to answer a particular scientific question. The characteristics of these persons (i.e., the "sampling frame") dictate both the methods that are applied to assess this population and how these methods are implemented. For example, if a particular research question pertains to the psychopathology among a small group of rescue personnel who assisted disaster victims, a list of these rescue personnel may exist, and it may be feasible to contact all these persons and interview them in person (by having interviewers visit them) or having them all come to be interviewed at a central location.

Telephone-based research methods, on the other hand, are particularly well suited to the sampling of large populations that are difficult to contact, or to interview, using other modalities. For example, a study that intends to sample all residents of a large city will have to devise a method to access persons who validly represent this sampling frame. In this context, problems about accessing buildings in dense urban areas and the size of the population to be sampled frequently make telephone-based recruitment and interviewing the only feasible option. Overall, investigator judgment that combines an awareness of the sample of interest, the key research questions and the assessment modalities that are needed to address them, and the practical and logistical (e.g., financial, time) parameters that bound the research will best guide choice of research method. In the rest of this chapter we will discuss considerations about the use of telephone-based research methods that may help inform such decisions.

Generating a Sample: List-Based Sampling, Random Sampling, and Oversampling

There are three principal methods of sample recruitment that are in use in telephone-based research, each with particular advantages and disadvantages. List-based samples involve the use of preexisting lists of all persons in a particular sampling frame. Such lists are useful when they represent the entire universe of persons in the sampling frame. For example, in research about mental health among families of the victims of the Pan Am 103 explosion over Lockerbie, researchers used a list of all family members, provided by official government sources, to contact all family members of

victims (Smith, Kilpatrick, Falsetti, & Best, 2002). Telephone interviewing then provides an opportunity to contact persons on these lists cost-effectively, particularly if these persons are geographically dispersed.

List-based sampling rapidly loses utility if the lists available are not comprehensive but rather include only a subgroup of all the persons in a particular sampling frame. For example, telephone directories provide a list of all persons within a particular city or region with listed telephone numbers. However, only about two-thirds of all U.S. households have listed telephone numbers (Survey Sampling Inc., 1990). In addition, the proportion of telephone numbers that are listed or unlisted in given areas throughout the United States varies, with a higher proportion of unlisted numbers in central cities than in suburban/rural areas. Therefore, if a telephone directory list is used as the primary means of characterizing a sampling frame, the sample will then best be construed as representative of *listed* telephone numbers. Epidemiologically, such samples are subject to *selection bias*. Given that we cannot adequately characterize the nature of this bias (e.g., are persons with listed telephone numbers younger or older than those with unlisted numbers?), the consequences of this bias are difficult if not impossible to assess, making it difficult to draw conclusions from the study. Therefore, list-based samples are useful only in a narrow subset of studies where a comprehensive list of persons in the sampling frame of interest is indeed available.

Random digit dialing (RDD), first proposed in the early 1960s, is a group of probability sampling techniques that provide an opportunity of reaching any household with a telephone within a sampling frame (Cooper, 1964). The primary advantage of RDD is that it removes the problem of selection bias arising from unlisted telephone numbers in specific areas. The science of generating random digit number lists is sophisticated, and a full discussion of the pros and cons of various techniques of generating random digit numbers is beyond the scope of this chapter (see Lavrakas, 1993). In brief, RDD surveying generally makes use of a sample frame that involves phone numbers that have been generated at random. These numbers include a combination of 3-digit area code and 3-digit prefix (sometimes called the telephone "exchange") that includes the area of interest for the survey (e.g., city or country), and a randomly generated series of 4-number suffixes. In the early inception of the U.S. phone systems 3-digit prefixes were tightly correlated with precise geographic areas. Therefore, given a particular geographic sampling frame selecting numbers within a given prefix made it possible to generate a random list of all numbers within a particular area. However, with the increasing mobility of telephone numbers, 3-digit prefixes are becoming less reliable in terms of identifying specific geographic areas. Increasingly, particular geographic areas are covered by multiple telephone prefixes, requiring that all prefixes be included in the

RDD sampling frame in order to ensure that all households in the frame have a nonzero probability of selection.

Generation of RDD lists for surveying includes simple random sampling (i.e., the generation of numbers within a given area code and prefix), modifications of a technique often referred to as the Mitofsky–Waksberg method, or list assisted sampling. The Mitofsky–Waksberg method uses two-stage sampling (Waksberg, 1978). In the first stage a telephone number is selected from a simple random list and dialed. If the number proves to be a residential number, the set of 100 numbers that include the same eight digits as the number dialed (called the primary sampling unit, or PSU) is retained for the second stage. Surveying then can be carried out within the second-stage PSUs, increasing the likelihood of working numbers and minimizing phone calls to nonworking numbers (Potthoff, 1994). List-assisted sampling further refines this method to cost-effectively sample number banks that contain the majority of working telephone numbers (Tucker, Lepkowski, & Pierkarski, 2002).

Although generally the goal is to sample subgroups of the population in proportion to their size, sometimes a small group, such as an ethnic group, is of particular interest for the research. In this case, the research design can involve *oversampling* of persons within that racial/ethnic group. This will then result in a greater number of persons within this group, allowing for more power for within- and between-group analyses. Such oversampling results in a sample with differential probability of selection, and such differential probability of selection must be accounted for by appropriate weighting.

Special Issues: Sampling Telephones versus Sampling People

Telephone-based sampling relies on sampling *telephones* rather than sampling people or households. There are three important considerations that arise from this. First, telephone-based research methods rely on persons having a telephone in their house. Therefore, by definition, persons who are homeless or who are living in institutions (e.g., prisons, long-term care facilities) are not part of the sampling frame. Another concern in this regard is the growing use of cellular telephones and, particularly, the proportion of persons who use only cellular telephones and do not have a home telephone. It is not current practice to include telephone numbers for cellular telephones in telephone sampling frames. The Telephone Consumer Protection Act of 1991 (TCPA) prohibits placing calls from automated dialing systems to numbers with services for which the called party is charged for the incoming call. Practically, since many cellular telephone calling plans still bill the plan holder for calls received (at least insofar as deducting airtime minutes), it is a potential violation of the law to include cell phone

numbers as part of a telephone-based sampling frame. In 2003, approximately 3% of persons in the U.S. used a cellular telephone as their only telephone, with up to 6% among certain groups, particularly persons under the age of 25 (Blumberg, Luke, & Cynamon, 2004). As the use of cellular telephones grows, newer techniques in telephone sampling may need to be developed.

Second, a household with more than one telephone has a higher likelihood of being sampled than a household with only one telephone. Statistical adjustment in the form of sample weights have to be applied to deal with this differential probability of selection.

Third, persons also have a probability of selection inverse to the number of people in the household. This is not an issue in telephone-based studies of entire households, but if the focus of the investigation is persons, this introduces an important differential probability of selection of persons living in different-sized households. For example, persons living in households with two persons have twice the likelihood of being sampled as persons living in households with four persons. There are two considerations that are relevant in this regard. First, once again, statistical adjustment for this differential probability of selection is needed. Second, it is not sufficient to sample whomever may answer the phone since this introduces selection bias (e.g., women may be more likely to answer the phone than are men). Rather, a system needs to be put in place to randomly select persons from a household for participation in the telephone-based assessment. The most commonly used methods in this regard are the Kish procedure, which asks interviewers to enumerate all eligible persons in a household and then uses an algorithm to randomly select a person in the household (Kish, 1949, 1965), and the birthday selection method, which asks for the person with the birthday closest to the interview date; both methods produce quasi-random assessments of respondents within a household (Groves & Lyberg, 1988).

CONDUCT OF TELEPHONE-BASED RESEARCH

General Principles regarding Telephone-Based Survey Development and Interviewing

The art of constructing good survey instruments has been the subject of several excellent reference texts (Dillman, 1978; Fowler, 1993). In planning surveys intended for telephone use, the principles that apply to the design of all good surveys apply here also. However, perhaps more than in-person interviews, telephone-based surveys need to be as simple as possible, at a sixth-grade reading level at the most, relying on straightforward communication of question concepts that can be effective on the telephone. Given the absence of face-to-face contact, the interviewer cannot rely on visual

aids, gestures, or demonstrations to explain a particular question, hence making clarity of questions paramount. In our experience, interviews that are designed with closed-ended questions are more effective for telephone-based assessments than are open-ended questions. It is simpler and ideal if respondents can respond to questions with a simple "yes" or "no," a number (as in age), or one-word answers or phrases that are selected from rating scales or fixed-answer alternatives. In addition to minimizing mistakes in responses and inter-interview variability, such question wording increases the likelihood that participant responses are confidential since, even if someone were listening to the respondents' answers, they would be unable to determine the meaning of the responses.

Keeping telephone-based surveys as brief as possible is critical to maximize participant survey completion. One of the principal contrasts between telephone-based surveying and in-person interviews is that the rapport between participant and interview is more limited in telephone-based surveying. In addition, the burden is on the interviewer to keep the participant on the telephone. If the survey becomes too burdensome, it is easy for the participant to disconnect the phone and end the survey. Therefore, surveys must encourage participant engagement and minimize participant fatigue. In our experience, more than 95% of participants who start participation in postdisaster research will complete surveys that are approximately 30 minutes in length (Galea et al., 2004).

Conduct of Surveys

Interviews must be designed in such a manner as to make participants comfortable with the interview process, to ensure that participants understand questions that they are being asked, and to allow them the option of stopping the interview at any time and requesting help if they need it, particularly in the immediate postdisaster context.

Upon contact with a potential respondent, telephone surveys typically include an opening script that is designed to (1) introduce the purpose of the call, (2) explain who is calling, (3) disclose how long the interview will take, (4) explain any relevant survey procedures, (5) notify participants of institutional review board oversight of the research and their rights as participants in research studies, and (6) offer a telephone number (ideally toll-free) that a participant can call to verify survey authenticity. Such opening scripts, carefully and concisely worded, are the first contact between an interviewer and a participant and are critical in establishing a rapport between the two that will serve the research well throughout the course of the interview. Verbal informed consent is typically obtained as part of the opening script. In some telephone-based studies, particularly ones that con-

sist of telephone-based interviewing of participants recruited off existing contact lists, written information about the study is sent to participants ahead of the actual telephone contact, facilitating both the opening script and the consent process. After the opening script is read, permission to proceed with the interview is obtained from all participants. It is helpful for interviewers to have available to them scripts for how to deal with frequently asked questions (e.g., "How did you get my number?"). All numbers in a telephone-based sampling frame are dialed 10–15 times at several different times of day in academic surveys (although frequently fewer numbers of times in market research) before being declared nonanswers.

Interviews are conducted as efficiently as possible, with the interviewers reading preset scripts and deviating as little as possible from the scripts and questions as written. Procedures should be in place for participant callback if participants need to stop during the call at any time. In the context of postdisaster research where there is a concern about participant distress, it is important to have mental health support systems in place that can provide assistance if participants are distressed and/or request such help. Generally, participation in trauma research is well tolerated, and participants report positive effects of participation that compensate for any modest levels of upset or distress experienced during research assessments (Newman & Kaloupek, 2004, p. 383; see also Fleischman, Collogan, & Tuma, Chapter 5, this volume). Given the greater anonymity afforded by telephone-based assessments, it is plausible that distress is further minimized using these techniques, although we are not aware of assessments that have compared distress after telephone and other forms of assessment after disasters. However, the presence of even a small number of respondents who may want mental health assistance suggests the need for a mental health backup system for research conducted soon after a mass trauma (Galea et al., 2002; Galea et al., in press).

One of the key questions that often confront researchers conducing postdisaster assessments is whether there is a need to provide a financial incentive to participants. Nominal financial incentives ($10–$25) probably do not substantially improve outcome rates for short cross-sectional surveys or limited assessments but may be helpful as part of a tiered approach to recruiting, and retaining, persons in a telephone-based cohort study (Bucuvalas, Morgan, & Galea, 2002). If a decision is made to provide an incentive, the respondent's address needs to be obtained, usually at the end of the survey, for mailing of the incentive. Safeguards must be put in place to ensure that addresses and other unique identifiers are kept in a separate database from the actual survey responses, thus maintaining participant anonymity. This, properly explained, helps to increase respondent confidence in the confidentiality of their responses.

Professional Survey Management

The conduct of a telephone-based research study, like all studies, requires daily management and careful supervision. In that regard one of the more common questions faced by academic researchers in considering this work is whether such work can or should be carried out by trained professional surveyors or by volunteer surveyors (frequently students or trainees). Although, in theory, a rigorously trained staff of surveyors could carry out a telephone-based research project optimally regardless of whether they are paid or unpaid, paid staff are clearly incentivized to do well and can be better monitored and corrected when deviating from standard procedure. In addition, the intricacies of day-to-day management of telephone-based research methods are copious and may be challenging for nonprofessionals or researchers who are not immersed in a high volume of telephone-based surveys on a day-to-day basis.

As such, many telephone-based research projects are conducted by professional market research firms with research experience who work in collaboration with researchers on devising a research plan, developing a survey, and training surveyors. The professional firms then typically supervise day-to-day conduct of the survey and provide data to the researchers for analysis. These collaborations can be productive and highly efficient in the postdisaster context, maximizing the skills of both researchers and survey professionals, both bringing invaluable insights that can improve the conduct of telephone-based research (Galea et al., 2002).

Regardless of the management brought to bear on any particular research project, all interviewers must be trained and adept at administering the interview and dealing with potential participant questions. If a particular survey is being conducted in more than one language, interviewers who are fluent in the languages of the survey need to be available and a system in place to make sure interviewers who can speak a particular language are available to handle respondents who may request an interview in that language. Data monitoring for inter-interviewer variability and for internal survey consistency is essential on a day-to-day basis. Typically a certain proportion of interviews are audiotaped or listened to by supervisors to ensure fidelity to study protocols and the written interview. Recruitment and refusal rates should be monitored by supervisors and the project director daily.

Most telephone surveys are currently conducted using computer-assisted telephone interviewing (CATI). CATI systems control the handling of the sample telephone numbers and program the survey questions to be presented to the interviewer on a computer screen with controls for skip patterns, item rotation, and acceptable response ranges. In this way CATI-coded questionnaires can restrict possible responses to questions (hence minimiz-

ing both erroneous answers and interviewer coding errors) and, in more sophisticated uses of CATI systems, include algorithms that carry out checks for internal consistency on an ongoing basis through the survey. CATI-coded surveys can also be particularly useful in special contexts, such as longitudinal cohort studies where the CATI program can embed information about the respondent obtained in previous survey waves that can inform question administration in an ongoing wave (Lavrakas, Settersten, & Maier, 1991). With the sophistication of computer-controlled scripts, CATI telephone interviews can also use complex skip logic and qualifications to route respondents through question series that would be very distracting for an interviewer to administer without the assistance of a computer. Cantor and Lynch (2000) identify several other advantages of using CATI for telephone interviews. The first is that it permits greater quality control over interviewer behavior because of the standardization of how questions are presented in sequence to the interviewer. A second advantage is that CATI interviews usually occur in a centralized facility, which permits direct real-time monitoring of interviewers. One study randomly assigned interviewers to either a CATI condition or a standard interview condition in which interviewers used a regular interview to collect information via telephone from the interviewer's home (Hubble & Wilder, 1988). Respondents in the CATI interview condition had substantially higher disclosure rates for violent and nonviolent crimes.

Calculation of Telephone Survey Outcome Rates

There are two primary considerations with regard to the calculation of outcome rates in telephone-based research methods. The first of these is the calculation of outcome rates, and the second pertains to the significance of these rates. Pinpointing an exact outcome rate in telephone surveys is often difficult due to the challenge in classifying indeterminate cases. An exhaustive list of potential methods for calculating response rates is provided by the American Association for Public Opinion Research (AAPOR, 2003).

Briefly, the all-encompassing term that should be used (but is seldom encountered in academic research) is *outcome rates*; this term includes several other rates, including response rates, cooperation rates, and refusal rates. *Response rates* are the number of completed interviews divided by the number of all possible interviews. *Cooperation rates* are the proportion of all cases interviewed divided by all eligible cases. *Refusal rates* are the proportion of all cases in which a respondent refuses to be interviewed or cuts off contact before some predetermined point in the survey that represents completion. There are several ways to calculate each of these outcome rates. There are six different types of response rates, depending primarily on how partial interviews are considered. Since all possible interviews are

included in the denominator in these calculations (including cases of unknown eligibility), response rates are the most conservative outcome rate that can be provided. Cooperation rates are outcome rates among those eligible and therefore are higher than response rates; there are two cooperation rates that generally can be calculated, depending primarily on assessment of eligibility. There are three types of refusal rates that differ primarily on how the disposition of cases of unknown eligibility is treated. It is advisable to report several outcome rates, together with careful explanation of how they were calculated, in academic research publications.

The second key question in this regard pertains to the significance of outcome rates and whether there is a relationship between outcome rates and telephone-based methods' ability to provide valid estimates of population-based prevalences or associations of interest. Some of the best evidence that can be applied to answer this question arises from large surveys that have been ongoing for several decades and across regions, hence allowing comparisons of survey validity with changing response rates. The Behavioral Risk Factor Surveillance Survey (BRFSS) conducted by the Centers for Disease Control and Prevention (CDC) is an annual telephone survey aimed at assessing a range of behavioral risk factors nationwide (CDC, 2005). The BRFSS response rates vary from year to year and from state to state, but the overall decrease in BRFSS response rates is well documented, with median response rates across states falling from a high of 71.4% in 1993 to a low of 48.9% in 2000. Surveys conducted by official organizations, such as health departments and the government, typically have higher response rates than research surveys conducted by academic or nonprofit institutions (Mariolis, 2001).

Importantly, recent analyses of BRFSS data have shown that for a range of response rates for telephone surveys between 30 and 70%, the response rates were at most weakly associated with bias (Mariolis, 2002). In one analysis, it was shown that although a larger difference in response rate was associated with a larger difference in estimates of cigarette smoking prevalence between the BRFSS and the in-person Current Population Survey (CPS), the effects were small (Mariolis, 2002). Further evidence in this regard comes from an analysis that was designed to test potential differences associated with different response rates obtained from identical surveys. Keeter, Miller, Kohut, Groves, and Presser (2000) compared data from two surveys with response rates of 61% and 36% and found very few significant differences across 91 comparisons. In the field of mental health, analysis of data from the National Comorbidity Survey showed that the impact on population prevalence estimates of response rates differing between 74 and 82% were small (Kessler, Little, & Groves, 1995).

There are two last considerations with respect to outcome rates. First, extreme effort to recruit more reluctant nonparticipants may itself intro-

duce bias, since the reluctantly recruited participants potentially have reasons for providing false or misleading responses. This concern has long been raised by experts on psychological measurement and response-bias scales have been developed to allow for adjustment for or exclusion of inaccurate responses from survey assessments (Drasgow, Levine, & McLaughlin, 1987; Reise & Due, 1991). Large national surveys that invest substantial effort in maximizing response rates also include measures such as extensive interviewer training to increase rapport with participants in order to attempt to minimize this concern (Kessler, Mroczek, & Belli, 1999). Second, overzealous efforts to increase response rates may well be unethical if they violate respondents' stated desire not to participate in a particular survey assessment.

Telephone Surveys and Sampling Weights

Survey weights ensure that each person in the sampling frame has an equal probability of selection. These weights are inverse to the number of telephones in a household (to account for the fact that households with more telephones are more likely to be selected) and proportional to the number of persons within a household (to account for the fact that a single person represents other noninterviewed persons in the same household).

Other forms of sample weights also can be developed. In telephone-based samples that are recruiting persons in the general population from a particular geographic area, poststratification weights are frequently developed to account for discrepancies between the sample as recruited and anticipated distributions of persons within specific strata as expected based on U.S. Census data. For example, if young whites are undersampled, a sample weight is developed that is inverse of the proportional undersampling of this subgroup. Caution must be exercised in developing such weights for several reasons. In small samples, subgroup estimates may rely on very small sample sizes, and any deviation from the anticipated census distribution may be a function of random variability more than of any systematic undersampling. Sample weights reduce computational efficiency in standard statistical programs and, as such, should be used judiciously and only when clearly indicated. A final practical note in this regard is that certain statistical programs are not capable of accurately dealing with complex survey weights in many of their advanced procedures. For example, standard statistical procedures in SAS 8.0 with weights applied result in overestimation of sample size and in artificially low estimates of standard errors. SUDAAN (Shah, Barnwell, & Bieler, 1997) and STATA 8.2 (2004) are both statistical software packages that adequately take into account weighted survey designs.

ADVANTAGES AND DISADVANTAGES OF TELEPHONE-BASED RESEARCH METHODS

Advantages of Telephone-Based Research Methods

The principal advantage afforded by telephone-based research methods is the opportunity for a cost-effective and valid means of assessing persons in a large sampling frame. Starting with cost issues, the principal factor driving costs in data collection is typically personnel costs, primarily in the form of pay for interviews and data collectors. In-person interviewing, widely considered the gold standard in psychiatric epidemiology because it allows the administration of clinician-administered structured clinical interviews, is also the most personnel-intensive, and therefore most expensive, form of data collection. Although costs for all surveys vary tremendously, given a multitude of factors, in our experience in-person interviews can be as much as five times as expensive as telephone surveys.

The costs involved in any research method are largely a function of the length of time it takes to collect data on any given participant. The time per participant also then affects how long it will take to complete data collection on a given desired sample size. In-person recruitment or interviewing requires interviewers to go to a participant's address (be it home or otherwise) and incur substantial travel time between interviews. Alternatively, in-person interviews that rely on participants coming to a central research site require substantial coordination on the part of the research team to ensure that participant transportation is facilitated to, and that participants arrive at, a designated research site. In telephone surveying, interviewers can move from one interview to the next without leaving a central telephone surveying facility. It is also worth noting that in many densely populated urban areas access to residential buildings is difficult and, absent official permission or fiat, makes the reliable sampling of these residences infeasible.

Therefore, when compared to in-person recruitment, telephone-based methods allow researchers to contact more persons in a particular time for less money. These particular advantages make phone research methods particularly suitable for the postdisaster context. In many instances after disasters, an early postdisaster assessment that can establish baseline mental health estimates is a critical part of the research. In addition, delays in obtaining funding and the relatively limited availability of short-term funds make cost considerations paramount. Therefore, a survey technique that can be implemented relatively quickly and cost-effectively can enable research to be conducted in the postdisaster context that may otherwise not be feasible to implement. Telephone interviews using CATI in a centralized facility also offer quality control over data collection that is not possible in in-person interviews.

It is worth noting that there are other methods of recruiting participants that provide alternatives to either in-person or telephone-based recruitment. The reader is referred to other chapters in this volume for coverage of some of these methods, including Web-based research (Schlenger & Silver, Chapter 8, this volume). The ultimate decision about which study design optimally may be employed to address a particular research question rests not only on evidence of a method's feasibility and costs but also on evidence that a particular method can assess mental health reliably and validly. In that regard, a substantial body of evidence suggests that telephone research methods are a reliable form of epidemiological assessment.

Validity of Telephone-Based Assessments as Compared to In-Person Assessments

One assessment conducted both in-person and telephone interview surveys simultaneously in the same area using the same interview schedule (Weeks, Kulka, Lessler, & Whitmore, 1983). Response rates for the telephone survey were lower than for the in-person surveys, and telephone respondents tended to be younger, better educated, and more likely to be white than in-person respondents. However, and most importantly, the same assessment showed that there were no substantial differences in the accuracy of self-reported conditions or in health utilization questions. In fact, the assessment showed that internal consistency between responses was higher in the telephone surveys than in the in-person surveys. This study also showed that telephone surveys were appreciably cheaper to conduct than in-person surveys. A similar study (Aneshensel, Frerichs, Clark, & Yokopenic, 1982) found no statistically significant differences between the two interview methods for overall assessment of health status, illnesses reported for the preceding 4 months, or reports of hospitalization. In a recent analysis comparing national estimates of data from the BRFSS and the National Health Interview Survey (NHIS; which obtains information on medical conditions and health risk factors via within-household in-person interviews), it was shown that BRFSS estimates were similar to NHIS estimates for 13 of 14 measures examined, suggesting that any effect of telephone versus in-person interview on the quality of the information obtained was negligible (Nelson, Powell-Griner, Town, & Kovar, 2003). Specific to mental health, several studies have shown that telephone assessment of Axis I disorders (including depression and anxiety disorders) produced nearly identical results to in-person assessments using a variety of instruments (Simon, Revicki, & vonKorff, 1993; Paulsen, Crowe, Noyes, & Pfohl, 1988).

Another recent study compared in-person and telephone interviews with respect to their ability to detect exposure to violent victimization and several DSM-IV disorders among a sample of older adults (ages 55–85)

who were randomly assigned to be interviewed using a highly structured interview either via telephone or in person (Acierno, Resnick, Kilpatrick, & Stark-Riemer, 2003). There were no significant differences associated with interview mode for prevalence of victimization experiences or prevalence of DSM-IV disorders.

Limitations of Telephone-Based Research Methods

The absence of in-person contact limits the potential for biometric assessment that may increasingly be the key to answering important questions regarding the biological underpinnings of the psychological consequences of disasters. However, it is possible to locate and interview respondents via telephone and then collect biological samples by other means.

The dynamics of telephone-based interviewing are also different than those inherent in in-person interviewing. As noted previously, the absence of visual cues, the ease with which participants may terminate a telephone-based survey, and the need to keep the survey brief are all potential limitations to use of the method. In addition, telephone survey outcome rates have been falling over the past several decades, and the introduction of new technologies, including caller identification, and cellular phones are new challenges that may reduce outcome rates even further in coming years. Although the best evidence suggests that there is little evidence of systematic bias inherent in lower response rates that are obtained through best survey practice, greater effort may need to be expended in telephone-based sampling to obtain satisfactory outcome rates in order to provide confidence that a given study sample adequately reflects the underlying sampling frame.

Problems can also arise in telephone sampling when the sampling frame of interest is restricted to particularly small geographic areas, as in sometimes the case in studies of localized disasters. Frequently it is difficult for persons to accurately identify whether they live within or outside a survey's geographic boundaries, and it is challenging for an interviewer on the telephone to narrow down such boundaries absent visual aids (e.g., maps) that are relatively easy to use in in-person interviews. Algorithms can be developed and implemented through CATI coding that offer respondents dichotomous choices about where they live anchored to familiar features of the local landscape (e.g., are you closer or further from the park than this street?) that can narrow down respondent location to ensure it is within a desired sampling frame. Such algorithms however are time-intensive to implement.

Ultimately, the reliance on functioning telephones for this mode of data collection may limit the use of telephone-based methods in particular postdisaster contexts. In disasters that result in widespread system disrup-

tion or massive relocation, telephone surveys soon after an event may not be an appropriate method of participant recruitment. Of course, the same factors are present and present serious challenges to door-to-door or mail samples in similar contexts. In our experience, in the United States, telephone service is generally rapidly restored, and even persons who are displaced have telephone numbers forwarded, enabling contact relatively soon after a disaster. However, in the global context, particularly outside the Western world, the utility of telephone-based methods remains limited to countries with reliable telephones that can access a suitably high proportion of the population.

FUTURE OF TELEPHONE-BASED RESEARCH METHODS

We conclude with a note about the future of telephone surveys. As discussed in this chapter, telephone-based survey methods offer particular advantages for researchers interested in the consequences of disasters, particularly in large general population sampling frames. However, these advantages are offset by disadvantages that to some extent limit the appli cability of these methods. We suggest that most of the promise of telephone-based methods lies in the innovative use of telephone-based methods and the combination of research methods to maximize the advantages of different techniques.

In terms of innovative use of telephone-based methods, the application of telephone-based methods to study designs such as longitudinal cohort studies or case–control studies can make substantial contributions to the postdisaster research literature. For example, there is a paucity of research that has used telephone-based recruitment of general population representative samples for the purposes of longitudinal cohort studies in the aftermath of disasters. As a result we have little empiric evidence to inform our understanding of the general population consequences of disasters. Similarly, telephone-based methods can be used to recruit population-based controls for case–control studies where cases are persons with psychopathology after disasters, be it posttraumatic stress disorder in the long term or distinct patterns of psychopathology. Such study designs, facilitated by telephone-based methods, can push the envelope in postdisaster research and address questions that currently remain unanswered.

Combinations of research methods may hold particular promise. For example, persons in the general population can be recruited using telephone-based methods and invited to participate in subsequent research stages that involve biometric or in-person assessments. This combination of population-representative sampling and the opportunity for in-depth or

biological assessments may allow the extension of biological research in the postdisaster context to population-based samples, an extension that currently remains highly unusual.

Finally, the proliferation of Web-based and electronic forms of participant communication also presents an opportunity for telephone-based methods. In this case electronic communication with participants may be used both to assist in recruitment and to obtain ancillary participant data. In addition, it may be feasible to assess participants by using telephone-based methods after provision of information through the Internet, allowing the researcher to provide large amounts of information but to preserve person-to-person questionnaire administration.

ACKNOWLEDGMENTS

Preparation of this chapter was supported in part by Grant Nos. R25 MH070552, R01 DA 017642, R01 MH 66081, and R01 MH 066391 from the National Institutes of Health.

REFERENCES

Acierno, R., Resnick, H., Kilpatrick, D., & Stark-Riemer, W. (2003). Assessing elder victimization. *Journal of Social Psychiatry and Psychiatric Epidemiology, 38*, 644–653.

American Association for Public Opinion Research Standard Definitions. (2003). *Final dispositions of case codes and outcome rates for surveys*. Retrieved February 11, 2005, from www.aapor.org/pdfs/newstandarddefinitions.pdf

Aneshensel, C. S., Frerichs, R. R., Clark, V., & Yokopenic, P. A. (1982). Telephone versus in-person surveys of community health status. *American Journal of Public Health*, 72(9), 1017–1021.

Blumberg, S. J., Luke, J. V., & Cynamon, M. L. (2004, May). *The impact of wireless substitution on random-digit-dialed health surveys*. Paper presented at American Association for Public Opinion Research Annual Conference, Phoenix.

Bucuvalas, M., Morgan, M., & Galea, S. (2002, May). *Tracking continuing psychological outcomes from the World Trade Center Disaster among New Yorkers*. Paper presented at the American Association of Public Opinion Research Annual Conference, Nashville, TN.

Cantor, D., & Lynch, J. P. (2000). Self-report surveys as measures of crime and criminal victimization. *Criminal Justice 4*, 85–138.

Centers for Disease Control and Prevention. (2005). *Behavioral Risk Factor Surveillance System*. Retrieved February 2, 2005, from www.cdc.gov/brfss/

Cooper, S. L. (1964). Random sampling by telephone: An improved method. *Journal of Marketing Research, 1*(4), 45–48.

Dillman, D. A. (1978). *Mail and telephone surveys: The total design method*. New York: Wiley.

Drasgow, F., Levine, M. V., & McLaughlin, M. E. (1987). Detecting inappropriate test

scores with optimal and practical appropriateness indices. *Applied Psychological Measurement, 11*, 59–79.

Fowler, F. J., Jr. (1993). *Survey research methods* (2nd ed.). Newbury Park, CA: Sage.

Galea, S., Nandi, A., Stuber, J., Gold, J., Bucuvalas, M., Rudenstine, S., et al. (2005) Public reactions to survey research after terrorist attacks. *Journal of Traumatic Stress, 18*(5), 461–465.

Galea, S., Vlahov, D., Resnick, H., Kilpatrick, D., Bucuvalas, M., & Morgan, M. (2002). An investigation of the psychological effects of the September 11th attacks on NYC: Developing and implementing research in the acute postdisaster period. *CNS Spectrums, 7*(8), 593–596.

Galea, S., Vlahov, D., Tracy, M., Hoover, D., Resnick, H., & Kilpatrick, D. G. (2004). Hispanic ethnicity and post-traumatic stress disorder after a disaster: Evidence from a general population's survey after September 11. *Annals of Epidemiology, 14*(8), 520–531.

Groves, R. M., & Lyberg, L. E. (1988). An overview of nonresponse issues in telephone surveys. In R. M. Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls, & J. Waksberg (Eds.). *Telephone survey methodology* (pp. 191–212). New York: Wiley.

Hubble, D., & Wilder, B. E. (1988). Preliminary results from the National Crime Survey CATI Experience. New Orleans, LA: Proceedings of the American Statistical Association, Survey Methods Section.

Keeter, S., Miller, C., Kohut, A., Groves, R. M., & Presser, S. (2000). Consequences of reducing nonresponse in a national telephone survey. *Public Opinion Quarterly, 64*(2), 125–148.

Kessler, R. C., Little, R. J. A., & Groves, R. M. (1995). Advances in strategies for minimizing and adjusting for survey nonresponse. *Epidemiologic Reviews, 17*(1), 192–204.

Kessler, R. C., Mroczek, D. K., & Belli, R. F. (1999). Retrospective adult assessment of childhood psychopathology. In D. Shaffer, C. P. Lucas, & J. E. Richters (Eds.), *Diagnostic assessment in child and adolescent psychopathology* (pp. 256–285). New York: Guilford Press.

Kish, L. (1949). A procedure for objective respondent selection within the household. *Journal American Statistical Association, 44*, 380–387.

Kish, L. (1965). *Survey sampling*. New York: Wiley.

Lavrakas, P. J. (1993). *Telephone research methods* (2nd ed.). Applied Social Research Methods Series, Vol. 7. Newbury Park, CA: Sage.

Lavrakas, P. J., Settersten, R. A., Jr., & Maier, R. A., Jr. (1991). RDD panel attrition in two local area surveys. *Survey Methodology, 17*, 143–152.

Mariolis, P. (2001, October). Data accuracy: How good are our usual indicators? In *Achieving data quality in a statistical agency: A methodological perspective.* Symposium conducted at the meeting of Statistics Canada, Ottawa, Canada.

Mariolis, P. (2002, May). *Response rates and data accuracy.* Paper presented at the American Association of Public Opinion Research, Nashville, TN.

Nelson, D. E., Powell-Griner, E., Town, M., & Kovar, M. G. (2003). A comparison of national estimates from the National Health Interview Survey and the Behavioral Risk Factor Surveillance System. *American Journal of Public Health, 93*(8), 1335–1341.

Newman, E., & Kaloupek, D. G. (2004). The risks and benefits of participating in trauma-focused research studies. *Journal of Traumatic Stress, 17*, 383–394.

Paulsen, A. S., Crowe, R. R., Noyes, R., & Pfohl, B. (1988). Reliability of the telephone interview in diagnosing anxiety disorders. *Archives of General Psychiatry, 45*, 62–63.

Potthoff, R. F. (1994). Telephone sampling in epidemiologic research: To reap the benefits, avoid the pitfalls. *American Journal of Epidemiology, 139*(10), 967–978.

Reise, S. P., & Due, A. M. (1991). The influence of test characteristics on the detection of aberrant response patterns. *Applied Psychology Measurement, 15*, 217–226.

Shah, B., Barnwell, B., & Bieler, G. (1997). *SUDAAN User's Manual, Release 7.5*. Research Triangle Park, NC: Research Triangle Institute.

Simon, G. E., Revicki, D., & vonKorff, M. (1993). Telephone assessment of depression severity. *Journal of Psychiatric Research, 27*(3), 247–252.

Smith, D. W., Kilpatrick, D. G., Falsetti, S. A., & Best, C. L. (2002). Postterrorism services for victims and surviving family members: Lessons from Pan Am 103. *Cognitive and Behavioral Practice, 9*, 280–286.

Stata/SE 8.2 for Windows. (2004). College Station, TX: StataCorp LP.

Survey Sampling Inc. (1990). *A survey researcher's view of the U.S.* Fairfield, CT: Survey Sampling.

Tucker, C., Lepkowski, J. M., & Pierkarski, L. (2002). The current efficiency of list-assisted telephone sampling designs. *Public Opinion Quarterly, 66*, 321–338.

Waksberg, J. (1978). Sampling methods for random digit dialing. *Journal of the American Statistical Association, 73*, 40–46.

Weeks, M. F., Kulka, R. A., Lessler, L. T., & Whitmore, R. W. (1983). Personal versus telephone surveys for collecting household health data at the local level. *American Journal Public Health, 93*(12), 1389–1394.