This paper was submitted to the National Science Foundation as part of its SBE 2020 planning activity (www.nsf.gov/sbe/sbe_2020/). Its inclusion does not constitute approval of the content by NSF or the US Government. The opinions and views expressed herein are those of the author(s) and do not necessarily reflect those of the NSF or the US Government.

The Emerging Crisis in Sampling of Household Populations
A Challenge for NSF's SBE Directorate

By

Thomas M. Guterbock
Professor of Sociology
Director, Center for Survey Research
University of Virginia
Charlottesville, VA 22904-4767
TomG@virginia.edu

John Tarnai
Director
Social & Economic Sciences Research Center
Washington State University
Pullman, WA 99164-4014
Tarnai@wsu.edu

Abstract

It is rapidly becoming impractical to achieve, at reasonable cost, fully representative samples of the general population, whether at the local or national level in the United States. Face-to-face surveys using area-probability sampling on a national scale are far beyond the financial reach of most researchers or the capabilities of most survey organizations. The increasing difficulties and costs of gathering data from a representative sample of some known population threatens the external validity of sample surveys across the social sciences, whether for basic or applied research, and threatens NSF's and many other federal research budgets and the policies that rely on survey results. New methods of sampling are being devised and used, but none is thus far able to offer near-complete coverage along with acceptable rates of response at a reasonable cost. Resolving this coming crisis will require new methodological and statistical research that exploits newly available communication and database technologies and questions some of the fundamental assumptions that are commonly accepted in survey sampling. This paper is a call for a new, focused emphasis from NSF's Social, Behavioral and Economic Sciences Directorate on resolving the sampling crisis in these fields in the coming decade.

Statement of the Problem

The social landscape for survey research is changing dramatically as advances in communication technology continue. It is rapidly becoming impractical to achieve, at reasonable cost, fully representative samples of the general population, whether at the local or national level in the United States.

Face-to-face surveys based on area probability samples are still the gold standard. However, their cost is prohibitive and they are very time-consuming to deploy. Additionally, face-to-face surveys on a national scale are far beyond the financial reach of most researchers or the capabilities of most survey organizations.

This paper was submitted to the National Science Foundation as part of its SBE 2020 planning activity (www.nsf.gov/sbe/sbe_2020/). Its inclusion does not constitute approval of the content by NSF or the US Government. The opinions and views expressed herein are those of the author(s) and do not necessarily reflect those of the NSF or the US Government.

The problem of how to develop appropriate and affordable methods for selecting representative samples threatens the external validity of sample surveys across the social sciences, whether for basic or applied research. And the rising costs of achieving representative samples threaten NSF's and other federal research budgets and the policies that rely on survey results. This paper is a call for a new, focused emphasis from NSF's Social, Behavioral and Economic Sciences Directorate on resolving the sampling crisis in these fields in the coming decade.

The problem's importance

Survey research is a dominant activity in the social, behavioral, and economic sciences and is the main source of much of the data supporting research in these disciplines. Surveys are used by virtually all the federal statistical agencies. And scientific sampling is key to the research activities of most local and state government agencies and many firms in the private sector.

During the past 30 years Random Digit Dialing samples of landline telephones (landline RDD) has been the dominant sample frame because it ensured the most complete coverage of the household population and included unlisted landline households. Landline RDD telephone surveys are still widely used for surveys of adults and other household populations. But this approach is no longer a reliable means by itself for achieving a representative sample of the general population, as both coverage and response rates continue to change for the worse. Moreover, none of the various alternatives to date that have emerged is—at least as yet—a satisfactory solution.

Main Issues and Implications

Errors in 'representation' or 'non-measurement' in survey samples today stem primarily from coverage error and non-response error. These types of error have always been part of social surveys, but they are presenting greater difficulty today due to a variety of causes.

Response rates and contact rates have been falling and it is becoming increasingly difficult to reach people through traditional single mode survey methods. Households have various methods to avoid being contacted, including answering machines, voicemail, call blocking, caller-ID, etc. People are less cooperative than they were 30 years ago, and are more apt to simply refuse to participate in surveys.

Households are cancelling their landline telephone services and replacing them with cell phone services. Latest estimates from the CDC are that approximately one in four U.S. households now relies solely on cell phone telephone service. Although cell-phone RDD sample frames have become available in the past decade, their efficiency and geographic specificity are problematic, and they are very costly to implement.

Increasingly, individuals have their own cell phone and do not rely on a household landline for their communication needs. This opens new opportunities for contacting them outside of the home. For example, recent research with cell phone surveys of the general population in the U.S. shows a consistent pattern of about one in three of those who complete a cell phone survey do so away from home. Additionally, each member of a family may now have

This paper was submitted to the National Science Foundation as part of its SBE 2020 planning activity (www.nsf.gov/sbe/sbe_2020/). Its inclusion does not constitute approval of the content by NSF or the US Government. The opinions and views expressed herein are those of the author(s) and do not necessarily reflect those of the NSF or the US Government.

her/his own cell phone, so that households have multiple telephone numbers, calling into question the viability of the household as the unit of reference for the sampling frame. Multi-function portable and handheld devices, many with Internet capabilities, have been proliferating and continue to do so.

For local surveys, however, cell-phone RDD sampling frames may never be a cost-effective solution to the representation problem because of number portability. Now that individuals may keep a unique telephone number regardless of where they purchase the phone or choose to move, selecting a cell phone sample based on a geographic or political area is increasingly challenging.

One positive development for surveys is the advent of address-based sampling (ABS) which is based on the list of addresses used by the US Postal Service for delivery of mail. ABS offers virtually complete coverage of all U.S. households (approximately 98%). However, this sample frame does not include telephone numbers, and back-matching address lists to phone numbers generally succeeds for less than 50% of households. With its incomplete telephone coverage, ABS only partially solves some sampling problems, because in-person interviewing is extremely expensive, and mail surveys cannot be implemented quickly. Thus, the ABS method is only appropriate for those surveys with sufficient time to implement a sequence of mailings and allow returns to come in. Furthermore, mail questionnaires cannot be highly complex and they are subject to greater literacy bias than are interviewer-administered surveys.

A growing majority of U.S. households have access to the Internet and use it. Outside the home, many people have Internet access at work, at libraries, and at schools, and--using smart phones--many now have access essentially anywhere. Although the latest estimates (Spring 2009) from the PEW Research Center are that over 63% of households have Internet access in the home, there are no means for randomly sampling general population e-mail addresses. Attempts to use phone or mail to solicit general population responses by Internet have met with only limited success. Furthermore, the continuing 'digital divide' dictates that the poor, the elderly and minorities are underrepresented in the Internet frame, leading to non-negligible coverage error in many Internet surveys.

Increasingly, social media such as Facebook, Twitter, Linkedin, and others, are being used by individuals to connect and communicate with others. The implications for survey research of these social media have yet to be investigated

Because of nonresponse and coverage concerns, mixing survey modes is becoming a necessity for many surveys. Current thinking is that response rates and coverage will improve if people have multiple opportunities to respond by different survey modes. But as researchers rely increasingly on multiple ways for individuals to participate, there are negative implications for both costs and mode effects. The survey research literature documents the many different mode effects observed for telephone, mail, Internet, and face-to-face surveys. Combining survey modes in one survey project introduces mode effects into survey responses, and the field does not yet understand the implications of this for data quality and reliability.

This paper was submitted to the National Science Foundation as part of its SBE 2020 planning activity (www.nsf.gov/sbe/sbe_2020/). Its inclusion does not constitute approval of the content by NSF or the US Government. The opinions and views expressed herein are those of the author(s) and do not necessarily reflect those of the NSF or the US Government.

As more surveys migrate to self-administration via the Internet, mail questionnaires, portable devices and social media, visual design of the questionnaire becomes increasingly important. As Don Dillman and others have shown, the visual aspects of design significantly influence data quality and the rate of unit and item non-response to surveys.

In sum, apart from in-person interviewing with area probability or address-based design – both of which are extremely expensive and time consuming - no other survey modes provide full coverage along with good rates of response across social strata, age groupings and cultural groups.

Solutions being tested

A number of survey researchers, including Don Dillman, Robert Groves, Jon Krosnick, Paul Lavrakas, Michael Link, Doug Rivers, and others are exploring alternatives to traditional sampling methods. These include combinations of RDD and Internet respondents; dual-frame telephone samples combining RDD cell phone and RDD landline samples; triple-frame samples combining RDD cell phone, RDD landline and telephone directory samples; ABS samples matched back to telephone records; Internet panels recruited by RDD and then providing computers to randomly selected households, and various combinations of these. At the same time, Internet-based opt-in panels of respondents are now widely used by market researchers. Some of these approaches use non-probability methods of sampling. Current research is attempting to explore the viability and accuracy of these newer methods in dealing with coverage, nonresponse, sampling and measurement error. Despite the lack of data on the validity and reliability of some of these approaches, their use is growing rapidly: hence the need for accelerated and enhanced research on new methods and their outcomes.

What is needed?

We need new thinking that may question some longstanding survey assumptions and practices. Three key assumptions underlie the currently accepted approaches: 1) that we should reach general populations by approaching the household as a sampling unit; 2) that probability methods of selection must be used at every sampling stage; and 3) that current non-response errors and coverage errors are ignorable. But as we turn to modes and sampling frames that have clear coverage limits and low response rates, these assumptions are no longer tenable in many survey situations.

There are many questions about how to proceed. For example, is it better to continue to use households as the main sampling unit, or do we need to consider the individual as the main sampling unit? If cell phones can reach individuals in institutions, why are the residents routinely excluded from study populations? Are there acceptable ways to harness non-probability methods in the service of efficient, representative sampling that can pass scientific muster? Can non-probability samples be combined with probability samples? Can propensity weighting for opt-in panels be made to work in a reliable and valid manner? These are questions that survey researchers need to explore.

Solving these problems is important if we are to continue to be able to rely on surveys to produce reliable and representative data on household and other adult populations. The

This paper was submitted to the National Science Foundation as part of its SBE 2020 planning activity (www.nsf.gov/sbe/sbe_2020/). Its inclusion does not constitute approval of the content by NSF or the US Government. The opinions and views expressed herein are those of the author(s) and do not necessarily reflect those of the NSF or the US Government.

social science research funded by NSF is threatened if the costs of obtaining valid samples continues to skyrocket. Some of the major issues needing to be addressed include the following:

- What sample frames or combination of sample frames provide the most complete coverage and practical reachability of household and adult populations?

- How do individuals perceive their participation in surveys, and what are the factors that contribute to persuading them to participate and to having a good experience?

- What sequence of contacts by what survey modes is necessary to obtain results that are representative of the general population?

- What changes are possible in how we conduct interviews to both take advantage of the new social media, and to overcome the obstacles they present to contacting individuals? How, if at all, can social media be used to improve the contactability of survey samples, without jeopardizing privacy and confidentiality?

- Can propensity score weighting really solve some of the sample frame deficiencies of opt-in Internet panels?

- What visual design changes are beneficial to improved participation and data quality on mail and Internet surveys? Can Internet surveys take more complete advantage of the enormous flexibility of the computer and the ability to display visual elements as well as sound?

- How can we speed the move from 'tailored design' (which adapts the design for the needs of each study) to 'micro-tailored design,' which increases respondent choice and offers different protocols to different respondents in the same study rather than using a one-size-fits all approach?

Academic survey research centers, as they have in the past, will play a key role in experiments that will solve these problems, bringing together faculty and students with survey clients to identify new ways of conducting survey research. There should be partnerships with private-sector firms, especially smaller firms eligible for SBIR awards, which will mobilize the potential for innovation from the market research world and allow testing of some of their methods in a scientific research context. Full scale unconfounded experiments testing one sampling method against another can be expensive, but such experiments will be essential to the future validity of social science research.

This emerging crisis in sampling of general populations is being faced by all who gather and use social science data, and by every program in SBE. These programs must join with their research allies in AAPOR, AASRO, AStatA, and CASRO to give priority to resolving this emerging crisis in sampling of general populations in the coming decade.

This paper was submitted to the National Science Foundation as part of its SBE 2020 planning activity (www.nsf.gov/sbe/sbe_2020/). Its inclusion does not constitute approval of the content by NSF or the US Government. The opinions and views expressed herein are those of the author(s) and do not necessarily reflect those of the NSF or the US Government.

References

Dillman D. A., Smyth, J. D., & Christian, L. M. (2009) Internet, Mail, and Mixed-Mode Surveys (3rd Ed). New York: Wiley.

Lavrakas, P.L., and 20 others (2010). AAPOR Cell Phone Task Force 2010: New Guidelines and Considerations for Survey Researchers when Planning and Conducting RDD and Other Telephone Surveys in the U.S. with Respondents Reached via Cell Phone Numbers. To be posted at www.aapor.org on October 26, 2010.

Marsden, P. V. & Wright, J. D. (Eds.) (2010) Handbook of Survey Research (2nd Ed.). New York: Academic Press.

The authors wish to thank Don Dillman, John Kennedy, Paul Lavrakas, and Jim Wolf for their comments on an earlier draft.

This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.