1

Introduction

Humans have a long history of being interested in quantification, the process of representing something in some measurable (numerical) form. Quantities hold intrinsic value beyond their mere symbolic representations of amount. Although in earlier periods of history there was less need for the average person to have a broad understanding of quantity, that need now pervades our lives. Even persons who consider themselves nonquantitative cannot deny the myriad decisions they make daily based on considerations of quantity.

Since historical times it has come to be recognized that an exact count or measure is not always needed for effective decision making. This recognition has served as a basis for the development of sampling theory. Sampling, or the process of taking into account only a subset of all possible elements of a larger set or population of persons, places, things, and so on, has long been implicit in human judgment. (This serves as the underpinnings of stereotyping, for example.) Much more recent, however, has been the development of formal methods to engage in systematic sampling, which brings us to those techniques thought of as survey methods.

Survey methods are a collection of techniques for which the most typical purpose is to provide precise estimates (i.e., measures that are stable with relatively low variance) of the prevalence (i.e., amount) of some variable of interest: For example, what percentage of registered voters are likely to vote for a certain candidate; what percentage of households own videocassette recorders; what percentage of automobiles leave the assembly line with structural defects? In these examples, one does not need an exhaustive measure of all possible voters, households, or automobiles to gain information on which to base an accurate judgment about who will win the election, the size of the market for videotapes, or the quality of workmanship at a particular assembly plant.

When survey methods are properly employed, the resulting estimates can be extremely precise representations of whatever is being measured. So precise are they that their margin of error is negligible, at least from the standpoint of effective decision making. Yet valid survey methods (i.e., likely to be accurate) constitute a relatively new and evolving body of knowledge, which may explain why they are so often misused and

misinterpreted. Furthermore, sampling error is only one of many possible sources of survey error.

TOTAL SURVEY ERROR

In his distinguished treatise on the various potential sources of error in survey research, Groves (1989) makes it explicit that in addition to considerations of sampling error, a careful survey researcher must attend to methods that control and/or measure the potential effects of coverage error, nonresponse error, and measurement error. Thus, the emphasis that often has been placed on concerns about sampling error alone is "incomplete and unsatisfactory" (Groves, 1989, p. 13). Together, all potential sources of imprecision and bias constitute *total survey error* (cf. Fowler, 1993; Fowler & Mangione, 1990) and each should be considered separately when planning, implementing, and interpreting a survey.

Sampling Error

Too often, those who assess the validity and applicability of survey findings attend primarily (or even exclusively) to the size of the sampling error, which is a function of the heterogeneity of what is being measured, the size of the sample, and the size of the population. This source of imprecision in a survey is associated with the fact that only a sample of all elements in a population is studied rather than a census in which all elements would be studied. Thus, for example, a random survey with a sample size of 500 measuring a phenomenon that is distributed approximately 50/50 in a large population (say > 10,000) will have a sampling error of approximately plus or minus four percentage points at the 95% level of confidence. However, this does not ensure that the survey will have measured the phenomenon within four percentage points of accuracy, because there are many other potential sources of bias and imprecision beyond sampling error. Collectively, these other sources are often referred to as *nonsampling error*.

Coverage Error

Before one should reach a conclusion about how accurately a survey's findings generalize beyond the sample on which the data were gathered, one must consider whether all elements in the population had at least some chance (probability) of being sampled. For example, in all telephone surveys of the general public, those households without a telephone and the homeless have *zero* probability of being sampled. Thus, all telephone surveys are subject to the potential effects of coverage error if one attempted to generalize the findings to the entire public. To the extent that the phenomenon being studied is correlated with coverage/noncoverage, the accuracy of the survey's findings will be lowered. In the case of telephone surveys, because those citizens who cannot be reached via telephone have, as a group, lower incomes than the population with telephones, any telephone survey is likely to find somewhat higher levels of income and income-related behaviors among its respondents than exists in the overall population.

Nonresponse Error

Few surveys achieve a 100% response rate; that is, most surveys sample elements (e.g., people or households) from which no data are gathered. (This is due to a variety of reasons, such as refusals, vacations, and so on, that are discussed throughout this book.) A survey's findings will be subject to nonresponse error to the extent that the elements that are sampled but for which no data are gathered systematically differ from those sampled for which data are gathered. For example, if young adult males disproportionately refused to participate in a survey and/or were less likely to be interviewed because they were rarely at home when interviewers called, then any measure that correlated with age and gender (e.g., going to bars or playing contact sports) would be a less accurate (i.e., biased) measure of the total population due to nonresponse error. However, if the data gathered in a survey are not correlated with whether or not a given type of person responded, then error (bias) due to nonresponse would not exist.

Measurement Error

Not all data that are recorded on a survey questionnaire are accurate measures of the phenomenon of interest. These inaccuracies may be due to errors associated with the questionnaire, the interviewers, the respondents, and/or the mode via which the data are gathered (cf. Biemer, Groves, Lyberg, Mathiowetz, & Sudman, 1991). For example, a survey question may be worded poorly or the questions may be ordered in a way that distorts (biases) the answers that are given; interviewers may behave in ways that bias which answers respondents give; respondents may be unwilling or

unable to accurately respond to a question; and/or, the mode used to gather data (e.g., in-person vs. telephone) may contribute to measurement error.

Survey Costs

As Groves (1989) explains, efforts to reduce and/or measure the potential effects of the various types of survey error have real cost implications. The reader should note the basic distinction between approaches intended to reduce potential errors versus approaches intended to measure their potential effects. It may be too expensive to implement procedures that may eliminate (or substantially reduce) a potential source of error, but it might be affordable to implement a procedure to measure its approximate size and, thus, take it into account when interpreting the survey's findings. (The more advanced reader and practitioner are encouraged to study and restudy Groves's [1989] challenging, but nonetheless excellent volume, *Survey Errors and Survey Costs*.)

For novice survey researchers these considerations can seem forbidding or even overwhelming. When faced with all the potential threats to a survey's validity some may throw up their hands and question the value of the entire survey enterprise. But to do so is to fail to remember the fact that highly accurate surveys are routinely conducted by experienced surveyors who exercise the necessary care.

The present text serves as an introduction to these considerations as they apply to telephone surveys. This discussion is not meant to lower the esteem that good surveys merit or to dissuade anyone from conducting a good survey. Rather, it is meant to alert the reader to the many challenges one faces in conducting a telephone survey that will be "accurate enough" for the purposes for which it is meant.

My message to the novice should be clear: Planning, implementing, and interpreting a survey that is likely to be accurate is a methodical and time-consuming process, but one well worth the effort.

TELEPHONE SURVEYING IN PERSPECTIVE

Surprising as it may appear to some, telephone survey methods have undergone serious development only in the last 25 years. Prior to that time the proportion of households in the United States with telephones was too low to justify the use of the telephone as a valid sampling medium. Once the proportion of U.S. households with telephones exceeded 90%, which