
**Surveys**

A survey is a design for collecting data based on questions to individuals through a standardized instrument (questionnaire). It can be used within the more general research designs following the quantitative standard (experimental, quasi-experimental, nonexperimental designs). A fuller definition is as follows:

A “survey” is a systematic method for gathering information from (a sample of) entities for the purpose of constructing quantitative descriptors (statistics) of the attributes of the larger population of which the entities are members. The word “systematic” is deliberate and meaningfully distinguishes surveys from other ways of gathering information. The phrase “(a sample of)” appears in the definition because sometimes surveys attempt to measure everyone in a population and sometimes just a sample. (Groves et al. 2004, 2)

An example that may well depict this definition is the Survey of Consumers that has been carried out by the Survey Research Center at the University of Michigan since 1946. Each month, a sample of about five hundred adults is interviewed by telephone on consumption attitudes and behavior. The survey’s main objectives are to measure changes in consumers’ attitudes and expectations, understand why such changes occur, and evaluate how they relate to consumers’ decisions to save, borrow, or make discretionary changes. A term similar to survey is *opinion poll*: some authors use these terms interchangeably, whereas others believe that there is a difference between the two. According to the latter, the opinion poll is intended to answer questions (coming from a newspaper, a political party, a company, etc.) through a simple analysis of data collected on the basis of a short interview to a sample of individuals; in contrast, a survey aims to corroborate a more or less articulated theory, and complex analyses are developed on those data collected through extensive interviews.

**Data Collection**

There are three important aspects to consider when conducting a survey: the mode and the organization of the data collection, the sample design, and the construction of the questionnaire. Regarding the first aspect, it is possible to identify four major modes of collection: face-to-face interviews, telephone interviews, self-completed questionnaires, and self-completed questionnaires via the Internet. This typology is essentially based on two dimensions: the presence or absence of an interviewer and the type of support (via telephone or not in the case of the interview, through the web or not in the case of self-completed questionnaire). Currently, researchers use each of these different ways of collection, though in different proportions. However, they have characterized different phases in the history of the survey, reflecting the development of sampling techniques on the one hand and technology on the other. Whereas initially many individuals were interviewed through self-completed questionnaires, mainly via postal service, developments in sampling theory have reduced the number of individuals to be interviewed but have increased the quality of the collection process. The subsequent history of survey was instead marked by technological developments: From the early 1980s, the spread of the telephone, with its advantages in terms of savings and lower invasiveness, has led to an exponential growth in the use of telephone interviews. Lately, the diffusion of the web has revived the self-completed questionnaires. Finally, almost all of these modes of data collections are increasingly computer-aided: the support of the personal computer (PC) is obvious to the self-completed questionnaires via the Internet, but it is pervasive also for telephone interviewing and it is becoming more and more widespread in face-to-face interviews. One of the advantages of using computers is their assistance in running the sequence of questions and checking the compatibility between the codes keyed in and those expected under the coding scheme of responses, so completing the questionnaire becomes faster and the chance of errors decrease.

The four modes of data collection described in the previous paragraph can be distinguished
according to three dimensions: the quality of the process of administration, coverage and nonresponse errors, and organizational aspects. The following are general considerations that can result in significant exceptions in specific situations depending on the particular characteristics of the sample to be examined. Regarding the quality of the process of administration, the best case is basically face-to-face interviews where the interviewer can motivate the interviewee—the simple interaction creates favorable conditions for the involvement—and guide him or her in filling out long and complex questionnaires. In telephone interviews, the lack of personal contact may lower interviewer commitment and interviewee motivation, thus requiring some simplification of the questions; the interaction as well tends to lose importance more quickly, so the interview will generally be shorter unless the topic of the interview is particularly relevant to the interviewee. These disadvantages are increased when the interviewer is completely absent. Here we must distinguish between the classic self-completed questionnaires (such as postal ones) and self-administration that is still supported by an interviewer, as in group administration (e.g., school students) or on-site surveys (e.g., exhibition visitors). In the first case, the lack of control over the process of administration is complete, so there is no certainty about the identity of the respondent. In the second case, the support of an interviewer in some way mitigates what was said before, though the quality of the responses continues to depend on the motivation and cognitive skills of the respondents. Self-administration can be an advantage when the questions relate to sensitive issues (sexual behavior, drug use, etc.): in this case, in fact, the interaction with the interviewer may influence the answers of the interviewee.

The second dimension concerns the implications of different ways of collecting data on coverage and nonresponse errors. Face-to-face interviews, for example, require sample designs compatible with the spatial concentration of the sample in areas easily accessible to interviewers, but this requirement is largely irrelevant to telephone interviews, postal questionnaires, and those administered via the web. On the other hand, the particular characteristics of telephone (fixed and mobile) and Internet users—characteristics in rapid change in recent years—pose problems of coverage. In the case of fixed phones, there is a problem controlling the selection process of the respondent among the individuals present at home at the time of the interview: the underrepresentation of elderly and poorly educated people—which is common in telephone samples—is due to the tendency of these groups to leave the phone to other family members present in the house. The problem of self-selection—that only a few agree to the interview and the choice to respond is not random but associated with the particular characteristics of individuals—is pervasive in all modes of data collection but is particularly relevant in self-administered questionnaires. It can, however, be mitigated by the presence of an interviewer during the administration (as in group administration) or by an obligation to return the questionnaire, a situation that is rather rare. With regard to the self-administered questionnaire via the Internet specifically, at least two phenomena should be highlighted: on one hand, the use of the tool is rapidly increasing; the other is the risk that the share of so-called professional respondents becomes nonnegligible. These are individuals who have a stable relationship with research institutes and are rewarded in some way for their participation in the various surveys: these professional respondents could simply give “sensible” answers just to continue their participation, rather than in a manner corresponding to their own feelings, as is probable with those who are sufficiently motivated to answer without any reward other than the reaffirmation of their identity.

Finally, the third dimension concerns organizational aspects. From an economic standpoint, the most expensive mode of collection is the face-to-face interview, followed by telephone interview, while the self-completed questionnaire (via the Internet or not) is usually the cheapest. From a time perspective, the fastest ways for collecting data are via telephone and the web: this explains their intense use in opinion polls for the speed with which one can get the results that can make an important point for polls purchasers. An important organizational aspect of telephone surveys is the concentration of interviewers in one place: this facilitates not only their preparation and supervision but also group cohesion. This may be important for improving the environment and thus the quality of the work of interviewers. Whatever data collection mode is chosen, pretest is a crucial factor for survey success. Pretesting gives the research group an opportunity to test every single question and the questionnaire as a whole; it is also useful for
evaluating the duration of the interview. Although this entry has discussed distinct modes of data collection, a combination of modes can be used within the same survey, depending on the specific objectives and organizational constraints.

Sample Design

When researchers choose the collection mode to adopt in a survey, coverage and nonresponse errors are an issue; when they choose the sample design, sampling error, another important source of uncertainty about data, is a concern. Often, the research question aims at studying a big population, but it is simply impossible to survey all of the individuals belonging to it. So it becomes necessary to draw a sample from that population in a way that the small sample will make a reliable representation of the big population. Even if the population is not so big, it may be convenient to sample the population: the resources (money, time, organizational burden) saved on quantity (interviewees) can be spent on quality (of data collection). There are two kinds of sample designs: if units have a known (and different from zero) chance of selection, the sample design is probabilistic; otherwise, the sample design is nonprobabilistic. In other words, a sample design is probabilistic if John (and every other population unit) has a chance to be interviewed and this probability is known by the researchers.

Simple random sampling is the basic probabilistic design. It requires the full list of the members of the population that the researchers are studying: a number is assigned to every unit in the list and then numbers are randomly extracted. This way any population unit has a probability to be sampled, and that probability is the same for all members. Although simple random sampling is not complex as far as technical procedure and statistical theory are concerned, it is employed less than expected. If the population is big, it is difficult to get the full list of the members; moreover, if the researchers plan to use face-to-face interviews, it would be impossible to reach a population scattered everywhere. A version of simple random sampling is systematic sampling, which is different only with regard to the technical procedure of extraction. Units are in fact extracted according to a sampling interval that equals the size of the population divided by the size of the sample \( k = \frac{N}{n} \). For example, if the researchers want to draw a sample of 500 cases out of a population of 8,000, they have to extract a unit every 16 units (starting from a random number between 1 and 16).

Systematic sampling is particularly useful since it can be employed even if there is no list of the population; for example, if the researchers want to interview the customers of a supermarket, they could ask to interview every \( k \)th (according to the sampling interval) customer coming out of the store. Stratified sampling takes advantage of the fact that accuracy in sample estimates depends not only on sample size but also on the variance of the phenomenon under study. Population is divided into strata according to one or more variables connected to it (e.g., if the researchers want to estimate income, they could divide the population according the profession); then, inside each stratum a random sampling takes place. The final sample is smaller than a standard simple random sample, and the estimates are as accurate, if not more. Notwithstanding a complex theory and some statistical disadvantages, multistage sampling could be a suitable design when there is no list of the members of the population, when the population is scattered everywhere, or when it is necessary to save on data collection costs. In this design, the population is divided in hierarchical levels and at every level an extraction is carried out; for example, if the researchers want to study the workers of a big chain of supermarkets, they could sample first supermarkets and then workers.

Moving to nonprobability sampling, the most discussed design is quota sampling. The population is divided into strata, or quotas, like in stratified sampling, but with the important difference that inside each quota the selection of the individuals is not random but delegated to the interviewer. Such a nonrandom mechanism has been highly criticized by most statisticians as a source of bias because interviewers may tend to select individuals easier to interview; some survey researchers confute it, maintaining that practical advantages overcome statistical disadvantages. In the case of small populations, with individuals scattered everywhere but in contact with one another—as is often the case of “underground populations” (illegal immigrants, homosexuals, drug addict individuals, etc.)—a useful strategy could be that of snowball sampling. The researchers start with individuals selected in a particular way and ask them to indicate other members of the population, and so on: this way there is an exponential growing of the
sample. The risk of this design is that the snowball may take a particular direction, thus resulting in the sample not being representative of the surveyed population.

**Questionnaire Definition**

The last important aspect to consider when conducting a survey is the definition of the questionnaire. The questionnaire, compared to the questions schemes used in other research designs, is so characterized: the definition of the issues to deal with in the interview is specific; questions are the same in the same order to all the respondents; and most of the time, the respondent is not free to express himself or herself but has to choose among alternatives. The questionnaire is usually structured in three sections: initial, central, final. The initial section is the most important because the researcher has to persuade the sampled individual to give the interview: the potential respondent has to be reassured about survey objectives and purchaser, anonymity, and other aspects. Standard sociological questions (gender, age, etc.) can be put at the beginning of the questionnaire to be sure this information is obtained; however, if the researcher is concerned that the respondent could get bored or upset, standard sociological questions can be moved to the final section. Difficult questions (related to delicate and scabrous issues) can be asked in the central or in the final section: the researcher has to balance the importance of the questions and the risk of the interviewee quitting the interview, which would result in the loss of answers to the remaining questions. From a psychological point of view, the aim of the questionnaire is to activate “optimizing” rather than “satisficing” answer mechanisms: the respondent should give the best answer and not an acceptable answer. There are, in particular, two phenomena that can decrease answer quality: *social desiderability* (the tendency to introduce oneself in a favorable light) and *acquiescence* (the tendency to answer affirmatively regardless of the content of the question). To activate optimizing answer mechanisms and to avoid social desiderability, acquiescence, or both, question wording is extremely important. The usual norms apply: statements must be clear, concise, straightforward, and unambiguous; the vocabulary must be as simple as possible, according to the cognitive skills of the respondents; and so forth.

Surveys are rich sources of information. Once the survey project comes to an end, its data and results can continue to be useful: the survey can be repeated over time to study changes; data can be analyzed according to different research questions; results can be compared to reach stronger conclusions.

**Ferruccio Biolcati-Rinaldi**

*See also* Attitude Surveys; Likert Scales; Longitudinal Studies; Methodologies for Studying Consumer Culture; Methods of Market Research; Opinion Polls

**Further Readings**


**Symbolic Capital**

Symbolic capital is one of the forms of capital central to the work of Pierre Bourdieu. Although it is often simply glossed as “honor” or “prestige,” it is important to note that the honor and prestige inherent in...