Sheatsley, Paul B. "Questionnaire Construction and Item Writing." in *Handbook of Survey Research*. edited by Peter H. Rossi, James D. Wright, and Andy B. Anderson. New York: Academic Press, 1983, 195-230.

Chapter 6

Questionnaire Construction and Item Writing

Paul B. Sheatsley

6.1. INTRODUCTION

Although survey research implies almost by definition some sort of questionnaire to be administered to a sample of individuals, it is important to realize that the questionnaire is simply an instrument, a tool, to be employed in the study of a research problem. As such, it may or may not be the most suitable instrument for the task.

Some or much of the information to be sought may already be available in, for example, statistics compiled by federal government agencies or other sources, or in the files of such survey research archives as the Roper Public Opinion Research Center or the Interuniversity Consortium for Political and Social Research. For some studies, direct observation or measurement, such as counting traffic or the Nielsen television ratings based on actual operation of the television set, may be superior to retrospective questioning of individual respondents. Field experiments, in which the investigator devises a scenario and then records people's responses to the contrived situation, have a long history in social research. Content analysis of newspaper or magazine articles and advertisements may provide a better record of changes over time than asking people today about their recollections of the past. All of these methods have their own weaknesses, of course, as does using questionnaires; the point is that we should not automatically assume that a brand new questionnaire is the only means to provide the answer to every research problem.

If it appears that a questionnaire is the most suitable instrument for the research task, one must still ask whether it can do the job. It may be, for example, that people simply do not have the information: They cannot recall it,

HANDBOOK OF SURVEY RESEARCH

ISBN 0-12-598226-7

cannot predict it, or are ignorant of the matter under investigation. The problem of ignorance can be controlled, of course, by careful sample selection. Obviously, one would not ask the general public about the editorial policy or content of the *American Sociological Review*. Such a survey should be restricted to readers of the journal or to a similar professional audience. Yet most of us have been exposed at one time or another to interviews or questionnaires that have sought our opinions about subjects far removed from our ken. We may also have been asked to provide information on our job experience or salary levels over a long period of time, our childhood medical history, our expectations or motivations in some long ago period, or other details of the past that are either beyond our recall or subject to wild error. Similarly, questions that ask us what we would do if we inherited $1 million or what we expect to be the biggest problems facing us 10 years from now, or even what our plans are for the next year or two, can elicit little more than "best guesses" with a wide margin of error. A questionnaire can bring out only what is in the mind of the respondent, and this is task enough. To probe for what is not in the respondent's mind is a disservice to the respondent and is also unlikely to produce data of any validity and utility.

One other consideration before deciding to use a questionnaire is the willingness or readiness of the respondent to reply. Certainly the best questionnaire writers today, aided by changes in the social climate, have pushed back the frontiers from the recent past when it was generally assumed that one could not ask questions about drinking, drugs, sexual behavior, income, cancer, and a host of other taboo items in a household interview situation. But even today valid answers to such topics require careful introduction, proper survey auspices, and a well-planned line of questioning that does not depend on one or two blunt items. Even today, one may well question whether the questionnaire is the most suitable means of collecting data on such phenomena as the incidence of child abuse or the proportion of people who indulge in various forms of cheating, or even of measuring the extent of such socially disapproved behavior as gambling, "deviant" sex practices, or excessive use of alcohol or drugs. Obviously, any questionnaire, even if it consists of only a single question on the topic, can produce some data. In general, the more questions one asks on the topic, the more money one is prepared to spend on the survey, and the more proficient the line of questioning, the better the data will be. But there is always a point at which further investment will fail to produce equivalent returns—and sometimes this point is reached very early. Sometimes a little thought will show that a questionnaire is simply not worth the available time and budget, because people either do not have or will not easily reveal the information that the researcher seeks.

6.2. STANDARDIZED QUESTIONNAIRES

Once researchers decide that questioning people is the best way to obtain the necessary information, they are then faced with the task of designing a

suitable instrument. If researchers themselves or a small team of colleagues are collecting the data, and if the data are not to be handled statistically, it may be that no formal questionnaire is required. In interviewing community leaders, for example, about some topic of local concern, a standardized questionnaire may inappropriately narrow the discussion and prevent a full exploration of each respondent's views. Instead, one might prepare a brief interviewer guide, listing perhaps a dozen major questions, with appropriate probes listed under each. This will ensure that all obvious items are covered, but will allow ample room for the researchers to probe the unexpected response or to follow up on unforeseen factors they may have overlooked. Such interviews can be tape recorded, classified crudely by a number of dimensions, and analyzed qualitatively rather than statistically.

But if the researcher requires a large sample, numbering in the hundreds or the thousands, if the services of many interviewers will be employed; and if the data will be subject to statistical analysis, the task of designing a standardized instrument cannot be avoided. There must be a prescribed wording for each question, so that each respondent receives the same stimulus. One cannot have different interviewers asking the various items the way it seems best to them or improvising variations on the wording for different respondents. There is abundant evidence that even slight variations in question wording can significantly affect response (Payne, 1951). There must also be a prescribed order for asking the questions, and for the same reasons. Again, there is evidence that responses to certain kinds of questions vary significantly depending on the items that precede and follow them (Bradburn & Mason, 1964). Finally, there must be prescribed definitions or explanations to ensure that the questions are handled consistently. If one asks, for example, *Do you read a newspaper regularly?* the word *regularly* (and perhaps the words *read* and *newspaper* as well) must be defined for the interviewers, so they can answer respondents' requests for clarification if they occur, and so they will know how to probe such replies as *Regularly every Sunday* or *I read it more often than not*.

There are obvious disadvantages in using a standardized questionnaire, and one hears them often from critics of survey research: People understand the questions differently; respondents are forced into what may seem to them an unnatural reply; they have no opportunity to qualify their answers or to explain their opinions more precisely; they may feel they have already answered the question when the interviewer asks other prescribed questions on the same topic; and so on. But there is really no alternative to the use of standardized questionnaires in large-scale surveys. Without standardized question wordings and sequence, and standardized instructions to interviewers, researchers would be unable to measure or control response effects; they would receive an unacceptable number of uncodable responses; and they would be completely overwhelmed by the sheer mass of idiosyncratic material. Thus arises the challenge to the researcher—to design and develop a standardized instrument that will meet the data needs, but is also crafted well enough that every respondent will grasp the intent of each item and will be willing and able to respond to it.

An important consequence of the use of a standardized questionnaire is that, once it is printed, the researcher is committed to it and can do little or nothing to improve it. In informal interviewing, one can adapt questions, shift tactics, change procedures as the data collection proceeds. With a standardized questionnaire, administered by dozens of interviewers working simultaneously in many different locations, the researchers may not even be aware of any problem with the instrument until the field work is largely completed. If they suddenly have some great new thought and now see that an important question has been omitted, they cannot suddenly stop 50 or 100 interviewers and tell them to add the item. These considerations emphasize the importance of careful design and pretesting of the survey instrument. The most ingenious sample design, skilled interviewing, and sophisticated analytical techniques cannot redeem a survey that asked the wrong questions or asked them poorly. For this reason alone, it is clear that questionnaire design is a crucial element in survey research.

6.3. MODE OF ADMINISTRATION

Many issues of questionnaire design hinge upon the mode of its administration. The two basic modes are self-administration by the respondent and administration by an interviewer who asks the questions. Each of these can be subdivided or combined in various ways.

In the most frequent type of self-administered questionnaire, no interviewer is present. The most common example is the mail survey (see Chapter 10 by Dillman). Respondents receive the questionnaire in the mail, read the accompanying letter or instructions, and fill it out at home or elsewhere at their leisure. The major advantage of this mode of administration is its low cost, since there are no interviewer time commitments or travel charges. Mail surveys may sometimes produce more valid responses to certain types of questions where the presence of an interviewer might be inhibiting (see Chapter 9 by Weinberg). They are often a good means of collecting data from very specialized and highly motivated groups, such as opera-goers or members of a professional organization. The disadvantages of mail surveys are many, and they often outweigh the advantages. Response rates are generally low, with resultant large biases. Less-educated persons may have trouble following the instructions. Inadequate answers cannot be probed for a more specific or relevant response. If respondents are puzzled by an item, there is no interviewer to explain it to them. Question order biases may also occur because the respondent can study the whole questionnaire before answering the first question. One is not always sure that the person to whom the questionnaire is addressed is the one who fills it out. If respondents are ignorant, they can look up the answer or ask someone.

Another type of self-administered questionnaire is frequently given to groups of respondents, with an interviewer present, such as students in a

classroom or a group of shoppers invited to a meeting room. Here the presence of the interviewer can ensure a high completion rate and can be used to reassure respondents, to answer their questions, and even to edit their completed questionnaires. Costs, of course, are higher when the group must be specially assembled and one or more interviewers are sent to the meetings. There is also potential bias if school children fill out questionnaires in the classroom, factory employees perform the task on company time, or shoppers are paid to come to the researcher's office. Furthermore, such samples are usually inefficient for large-scale research because members of the groups are not independently selected. One would have to sample a large number of high school classrooms, for example, to generalize about high school students as a group.

Personal interviews, in which an interviewer asks the questions and fills out the questionnaire, can similarly be subdivided into face-to-face interviews, usually in the respondent's home or office, and telephone interviews. Finally, the two major modes may themselves be combined, as in a personal interview which includes a self-administered component, or a self-administered group interview that has the interviewer later asking a series of questions.

The mode of administration is relevant to questionnaire design because a questionnaire designed for one mode may not be applicable to the other. An obvious distinction is that when the questionnaire is designed for personal interview, the interviewer who uses it will be a trained individual who is already able to or who can be taught to use it proficiently under all circumstances (see Chapter 10 by Dillman). The self-administered questionnaire, on the other hand, will be filled out by untrained respondents who never saw the instrument before, will never see it again, and have no particular incentive to try to do a good job. It follows that self-administered questionnaires should generally be kept as simple, short, and self-explanatory as possible. Instructions should be brief and clear, answer categories unambiguous, and the line of questioning should avoid complicated skip patterns. In contrast, questionnaires designed for personal interview need be limited in their length and complexity only by time and cost factors. Although not to be recommended to amateur or journeyman researchers operating with less than million dollar budgets, there are examples of surveys that have employed 150-page questionnaires of devilish complexity and detail and have achieved very high response rates.

6.4. TYPE OF SAMPLE TO BE INTERVIEWED

Besides the mode of administration—personal interview or self-administered—a second major variable affecting questionnaire design is the type of sample from whom data are to be sought. One can imagine, for example, various extreme types: on the one hand, elementary school children, high school dropouts, people with little understanding of the subjects being asked about; and on the other, professionals, graduate students, corporate officers, community officials. It is clear, whether the questionnaire is to be self-adminis-

tered or the questions asked by an interviewer, that the instrument design task will be quite different, depending on the sophistication of the group. In the first case, one would be dealing for the most part with subjects who have a limited attention span, perhaps little interest in the subject matter, and a low-level capability in handling abstract ideas and subtle differences. In the other case, researchers need not hesitate to probe their subjects in considerable detail, to introduce fine distinctions, and to ask the respondents to perform fairly complex tasks.

Since the general public consists by definition of *all* types of individuals—all ages, education levels, types of neighborhood, and life experiences—the researcher who wishes to study a population sample, rather than some specialized group of people, is faced with a difficult task. It is a test of the researcher's skill to design a questionnaire that is simple enough for universal comprehension but not so childish or elementary that it will serve to alienate the interested and well informed. The researcher must pay due heed to the complexity of the topic and show an understanding of all the relevant variables, but keep the language and sequence of the questions within the reach and comprehension of the least-educated member of the sample.

It might be noted that because questionnaires are usually written by educated persons who have a special interest in and understanding of the topic of their inquiry, and because these people usually consult with other educated and concerned persons, it is much more common for questionnaires to be overwritten, overcomplicated, and too demanding of the respondent than they are to be simpleminded, superficial, and not demanding enough.

6.5. QUALITIES OF A GOOD QUESTIONNAIRE

Unlike sampling and data processing, questionnaire design is not a science or technology but remains an art. Given the same research task and the same hypotheses, six qualified questionnaire writers will be likely to come up with six instruments that differ widely in their choice of items, line of questioning, use of open-ended questions, and length of time the interview takes. Frequently, a good a priori case can be made for any of them. Furthermore, all researchers know that when they start analyzing their data, they are sure to find that some of the questions are useless to the task, whereas others that are sorely needed were somehow omitted from the design. There are no pat or simple rules for questionnaire writing. Most texts and articles on the subject are pitched on a vague general level, such as "Decide what information you need," or they consist of highly specific admonitions, such as "Be sure alternatives are mutually exclusive." The authors of these instructions are experienced researchers who know their field, but it is very hard to tell someone how to design a useful questionnaire.

One might start by looking at the general purposes that any questionnaire is designed to serve. These seem in most cases to be three in number. A well-

designed questionnaire should: (*a*) meet the objectives of the research; (*b*) obtain the most complete and accurate information possible; and (*c*) do this within the limits of available time and resources.

To observe that a questionnaire should meet the research objectives may seem obvious, but it certainly does not happen every time, often because the questionnaire was poorly designed. Researchers become seduced by their own particular prejudices or interests into elaborating some less important variable while devoting less attention to a more important one. For lack of hard thought, consultation, pretesting, or familiarity with the literature, one may completely omit from the questionnaire some important aspect of the subject. Much of this is inevitable; none of us is omniscient. Indeed, the very fact that we are studying something means we do not know it all and thus cannot conceive of all possibilities. Every survey is bound to leave some questions unanswered and to provide a need for further research. But the purpose of a good questionnaire is to minimize these problems by trying to forestall them.

It is perhaps equally obvious that a good questionnaire should provide the most complete and accurate information possible. Even when the questionnaire is carefully crafted to meet the research objectives, the accuracy and completeness of the data it produces are far from guaranteed. This is because the people who provide the information—the respondents—intervene in the process. Respondents may misunderstand the questions, they may reject the premises on which the questions are based, or they simply may refuse to answer. Worse, they may lie to the interviewer or attempt to conceal their actual behavior or attitudes. The questions may be so far above their understanding and experience that they answer at random rather than confess their ignorance. When these things occur, the questionnaire is a poor one, no matter how well it seemed to meet the survey's objectives when it was designed. The good questionnaire must be organized and worded to encourage respondents to provide the most accurate and complete information they can.

The third property of a good questionnaire—to provide the required information within existing cost and time constraints—may not seem quite as obvious, but is nevertheless very real. A questionnaire that provides data too late to be of real use or which causes researchers to run out of money before they can complete the analysis is a bad questionnaire, no matter how well it seemed to serve the objectives of the inquiry. Clearly, if one had unlimited time and unlimited funds, and could spend hours and hours with the respondents, one could devise more informative questionnaires—although not necessarily more efficient ones.

But in the real world, time and cost constraints are always present. Of all the surveys conducted today, only a very few start with researchers figuring out what they want to do and applying for what they consider to be the necessary funds. Even in these cases, budgets are usually reduced before a researcher gets the grant. The typical case is that of a client or sponsoring agency who has a problem, who has a specified (and usually, to the researcher, inadequate) budget to support the research, and who asks the research[er] to design a

study that will fit that budget and can be completed by a particular deadline. There is nothing wrong with limited surveys conducted on modest budgets. Indeed, it would be hard to argue that the most expensive surveys and those that have taken the longest time to complete have produced the largest body of useful information or have contributed the most to the advancement of survey methodology. "Quickie" surveys conducted at limited cost often provide the most effective policy guidance.

Time and cost affect all aspects of survey research, of course, but they have special importance to questionnaire design because they impose a limit on the number of questions the researcher can ask. A crucial element in budgeting a survey is the average length of the interview because this affects the number of calls an interviewer can make in a day, the number of interviewers to be employed and trained, the amount of data that must be processed and analyzed, and so on. The length of the interview is usually predetermined—20 min, 40 min, 1 hr, or whatever—and the task of the questionnaire designer is to fashion the most useful instrument possible within that constraint.

6.6. DECIDING ON CONTENT

All questionnaire writers have their own approach to instrument design, and the various approaches will also be modified according to the objectives and circumstances of the research—whether writers craft their own studies or do a job for a client; whether they work alone, with colleagues, or as part of an organization; whether they seek descriptive data about some phenomenon or are testing research hypotheses; and whether they have much or little time and money available. The following five steps, however, are generally applicable.

1. Decide what information is required.
2. Draft some questions to elicit that information.
3. Put them into a meaningful order and format.
4. Pretest the result.
5. Go back to 1.

Note that one does not start by writing down questions. It is not difficult to think up a hundred questions about any problem worthy of study, and talking to other people will suggest hundreds more. But unless researchers have some conceptual or analytical framework to guide them, there is no particular reason for choosing any one question over another. Actual drafting of the questions can usually be done fairly quickly and there is no hurry about it. The first task, and the hardest, is to figure out which factors are relevant to the problem. What are the things one needs to know from the respondent in order to meet the survey's objectives? This is the first question that needs to be answered.

Most research begins with at least a vague notion of the kinds of information required, if only because someone must have been sufficiently impressed to budget money for the inquiry. If researchers themselves have submitted a

successful proposal, they have made a good start on the questionnaire design. If they are doing a job for a client, the client will have defined the problem and will have specified or suggested the necessary information. Researchers should know, however, that clients are very often vague or biased about their needs and it is the researcher's task to clarify the goals, try to establish operational measurements, and probe for other relevant variables that the client may have overlooked.

Additional help is available from many sources, of which one of the most obvious is the literature search. Researchers should know what work has been done on the same or similar problems in the past, what factors have not yet been investigated, and how the present survey can build on what has already been discovered. Another source of help is discussion of the problem with friends and colleagues. Conversations with them may suggest new ideas, alter one's assumptions, or help in clarifying doubtful points. Consultation with experts, people with firsthand knowledge of the problem or who have spent time studying it, are almost always helpful. These need not be paid consultants. Most scholars, businesspeople, government officials, and the like will be glad to spend half an hour or more discussing with a researcher their own ideas and experience about a topic that concerns them.

A final source of ideas in selecting the content of the questionnaire lies not in experts or in the literature, but among the population one intends to survey. A half-dozen exploratory conversations with representatives of the group, perhaps roughly stratified by age, sex, education, or other presumably significant variables, will often provide researchers with a glimpse of reality that may sharply alter some of their preconceptions. Researchers may find that some consideration thought to be important seems to be quite irrelevant to the people they are talking to, or that they have deep concerns that were not previously realized. Out of one or more of these kinds of efforts, researchers can usually develop a fairly good outline of the kinds of information they need to obtain from the survey instrument.

A Checklist of Variables

Researchers have long sought to codify the various dimensions of public opinion that should normally be covered in survey research. One of the earliest of these was Gallup's "quintamensional" technique, which he recommended as necessary for a full understanding of a person's opinion (Gallup, 1947).

First, he said, is the respondent's *knowledge* or awareness of the issue. Almost always, researchers should include some measure of knowledge, so they can distinguish among the well informed, the poorly informed, and the unaware. Knowledge is usually highly correlated with education, but not always. Knowledge can be measured by asking a single question or by asking a whole battery of questions. Second is the respondent's *interest* in the problem or concern about it. Some people are apathetic about the issue, others highly concerned. The researcher should be able to specify the general level of interest