
ASKING AND ANSWERING SURVEY QUESTIONS: COGNITIVE AND COMMUNICATIVE PROCESSES

Quite obviously, survey estimates are only as meaningful as the answers that respondents provide. In the total survey error framework the questionnaire is one of the sources of measurement error (deviations between answers of respondents and some average or "true" value of the measured attribute). Although this fact has been recognized since the early days of survey research (see Payne, 1951, for an early review), survey methodology has long been characterized by rigorous theories of sampling on the one hand, and the so-called "art of asking questions" on the other hand. Despite a tradition of carrying out split-sample experiments within surveys in order to address practical problems of question form, wording, and context (see Schuman & Presser, 1981, for reviews and extensions of much of this early experimentation), it has only been recently that the cognitive and communicative processes underlying question answering in surveys have received sustained theoretical attention from psychologists. Drawing on theories of language comprehension, memory, and judgment, psychologists and survey methodologists have begun to formulate explicit models of the question answering process and have tested these models in tightly controlled laboratory experiments and split-sample surveys. Several recent volumes reflect the progress made in this area (Hippler, Schwarz, & Sudman, 1987; Jabine et al., 1984; Jobe & Loftus, 1991; Schwarz & Sudman, 1992, 1994, 1996; Sudman, Bradburn, & Schwarz, 1996; Tanur, 1992). Below, we introduce respondents' tasks and subsequently address implications for behavioral reports and attitude measurement.

Respondents' Tasks

Answering a survey question requires that respondents perform several tasks, as first emphasized by Cannell and his colleagues (Cannell, Marquis, & Laurent, 1977). Subsequent researchers built on their model (Strack & Martin, 1987; Tourangeau, 1984; Tourangeau & Rasinski, 1988). Not surprisingly, the first task of respondents consists of interpreting the question to understand what is meant. If the question is an attitude question, they may either retrieve a previously formed attitude judgment from memory, or they may "compute" a judgment on the spot. To do the latter, they need to retrieve relevant information from memory to form a mental representation of the target that they are to evaluate. In most cases, they will also need to retrieve or construct some standard against which the target is evaluated.

If the question is a behavioral question, respondents need to recall or reconstruct relevant instances of this behavior from memory. If the question specifies a reference period (such as "last week" or "last month"), they must also determine if these instances occurred during this reference period or not. Similarly, if the question refers to their "usual" behavior, respondents have to determine if the recalled or reconstructed instances are reasonably representative or if they reflect a deviation from their usual behavior. If they cannot recall or reconstruct specific instances of the behavior, or are not sufficiently motivated to engage in this effort, respondents may rely on their general knowledge or other salient information to compute an estimate.

Once a "private" judgment is formed in respondents' minds, they have to communicate it to the researcher. To do so, they may need to format their judgment to fit the response alternatives provided as part of the question, and in many instances the offered alternatives may both suggest and constrain what is retrieved from memory. Moreover, respondents may wish to edit their response before they communicate it, due to influences of social desirability and situational adequacy.

Accordingly, interpreting the question, generating an opinion or a representation of the relevant behavior, formatting the response, and editing the answer are the main psychological components of a process that starts with respondents' exposure to a survey question and ends with their overt report.

Question Comprehension The key issue at the question comprehension stage is whether the respondent's understanding of the question does or does not match what the
researcher had in mind: Is the attitude object, or the behavior, that the respondent identifies as the target of the question the one that the researcher intended? Does the respondent’s understanding tap the same facet of the issue and the same evaluative dimension?

From a psychological point of view, question comprehension reflects the operation of two intertwined processes (see Clark & Schober, 1992; Strack, 1994; Strack & Schwarz, 1992). The first refers to the semantic understanding of the utterance. Comprehending the literal meaning of a sentence involves the identification of words, the recall of lexical information from semantic memory, and the construction of a meaning of the utterance, which is constrained by its context (see Mitchell, 1994; Simpson, 1994, for reviews). Not surprisingly, survey textbooks urge researchers to write simple questions and to avoid unfamiliar or ambiguous terms. Sudman and Bradburn’s (1983) Asking Questions provides much useful advice in this regard.

However, understanding the words is not sufficient to answer a question. For example, if respondents are asked, “What have you done today?”, they are likely to understand the meaning of the words. Yet, they still need to determine what kind of activities the researcher is interested in. Should they report, for example, that they took a shower? Hence, understanding a question in a way that allows an appropriate answer requires not only an understanding of the literal meaning of the question but involves inferences about the questioner’s intention to determine the pragmatic meaning of the question.

To understand how respondents infer the intended meaning of a question, we need to consider the assumptions that govern the conduct of conversation in everyday life. These tacit assumptions were systematically described by Paul Grice (1975), a philosopher of language (see Clark & Schober, 1992; Schwarz, 1994; Strack, 1994; Strack & Schwarz, 1992, for applications to survey research and experiments). According to Grice’s analysis, conversations proceed according to a cooperativeness principle, which can be expressed in the form of four maxims. A maxim of quality enjoins speakers not to say anything they believe to be false or lack adequate evidence for, and a maxim of relation enjoins speakers to make their contribution relevant to the aims of the ongoing conversation. In addition, a maxim of quantity requires speakers to make their contribution as informative as is required, but not more informative than is required, while a maxim of manner holds that the contribution should be clear rather than obscure, ambiguous, or wordy. In other words, speakers should try to be informative, truthful, relevant, and clear. As a result, “communicated information comes with a guarantee of relevance” (Sperber & Wilson, 1986, p. vi) and listeners interpret the speakers’ utterances “on the assumption that they are trying to live up to these ideals” (Clark & Clark, 1977, p. 122). These tacit assumptions have important implications for question wording.

Response Alternatives Suppose, for example, that respondents are asked in an open response format, “What have you done today?” To give a meaningful answer, respondents have to determine which activities may be of interest to the researcher. In an attempt to be informative, respondents are likely to omit activities that the researcher is obviously aware of (e.g., “I gave a survey interview”) or may take for granted anyway (e.g., “I took a shower”). If respondents were given a list of activities that included giving an interview and taking a shower, most respondents would endorse them. At the same time, however, such a list would reduce the likelihood that respondents report activities that are not represented on the list (see Schuman & Presser, 1981; Schwarz & Hippler, 1991, for reviews). Both of these question form effects reflect that response alternatives can clarify the intended meaning of a question, in the present example by specifying the activities the researcher is interested in. Moreover, the response alternatives may serve as reminders, bringing behaviors or opinions to mind that may otherwise not have been accessible. These processes result in pronounced differences between questions that use an open versus a closed response format (see Schuman & Presser, 1981, for examples).

Whereas the notion that response alternatives may clarify the meaning of a question may seem rather obvious in the above example, more subtle influences are often overlooked. Suppose that respondents are asked how frequently they felt “really irritated” recently. To answer this question, they again have to determine what the researcher means by “really irritated.” Does this term refer to major or to minor annoyances? To identify the intended meaning of the question, they may consult the response alternatives provided by the researcher. If the response alternatives present low frequency categories, e.g., ranging from “less than once a year” to “more than once a month,” respondents may conclude that the researcher has relatively rare events in mind and that the question cannot refer to minor irritations, which are likely to occur more often. Supporting this assumption, Schwarz, Strack, Müller, and Chassein (1988) observed that respondents reported examples of differentially extreme irritations, depending on the frequency range of the response alternatives presented to them.

Similarly, Schwarz et al. (1991) observed that respondents may use the specific numeric values provided as part of a rating scale to interpret the meaning of the scale’s labels. A representative sample of German adults was asked, “How successful would you say you have been in life?” This question was accompanied by an 11-point rating scale, ranging from “not at all successful” (combined with the numeric values 0 or −5) to “extremely successful” (10 or +5). The results showed a pronounced impact of the numeric values of the rating scale. Whereas 34 percent endorsed a value below the midpoint of the 0 to 10 scale, only 13 percent endorsed one of the formally equivalent
values on the -5 to +5 scale. Subsequent experiments indicated that this difference reflects differential interpretations of the term "not at all successful." When this label was combined with the numeric value of "0," respondents interpreted it to reflect the absence of success. However, when the same label was combined with the numeric value "-5," and the scale offered "0" as the midpoint, they interpreted it to reflect the presence of failure. In general, rating scales that use only positive numeric values suggest that the researcher has a unipolar dimension in mind (pertaining to the extremity of a single attribute, e.g., success), whereas a combination of negative and positive numeric values suggests that the researcher has a bipolar dimension in mind (pertaining to the presence of a given attribute or its opposite, e.g., success vs. failure).

Such findings demonstrate that respondents use the response alternatives in interpreting the meaning of a question. In doing so, they proceed on the tacit assumption that every contribution is relevant to the aims of the ongoing conversation. In the survey interview, as in laboratory experiments, these contributions include apparently formal features of questionnaire design, such as the numeric values given on a rating scale. Hence, identically worded questions may acquire different meanings, depending on the response alternatives by which they are accompanied (see Schwarz, 1996; Schwarz & Hippler, 1991, for more extended discussions).

**Question Context** Respondents' interpretation of a question's intended meaning is further influenced by the context in which the question is presented. Not surprisingly, this influence is more pronounced, the more ambiguous the wording of the question is. As an extreme case, consider research in which respondents are asked to report their opinion about a highly obscure—or even completely fictitious—issue, such as the "Agricultural Trade Act of 1978" (e.g., Bishop, Oldendick, & Tuchfarber, 1986; Schuman & Presser, 1981). Questions of this type reflect public opinion researchers' concern that the "fear of appearing uninformed" may induce "many respondents to conjure up opinions even when they had not given the particular issue any thought prior to the interview" (Erikson, Luttbeg, & Tedin, 1988, p. 44). To explore how meaningful respondents' answers are, survey researchers introduced questions about issues that don't exist. Presumably, respondents' willingness to report an opinion on a fictitious issue casts some doubt on the reports provided in survey interviews in general. In fact, about 30 percent of a sample do typically provide an answer to issues that are invented by the researcher, although the percentage of such answers can be reduced by including an explicit "Don't Know" alternative as part of the question, thus providing more legitimacy to the admission of ignorance (see Schuman & Presser, 1981).

From a conversational point of view, however, the answers provided to fictitious issues may be more meaningful than has often been assumed. The sheer fact that a question about some issue is asked presupposes that the issue exists—or else asking a question about it would violate every norm of conversational conduct. Respondents, however, have no reason to assume that the researcher would ask a meaningless question and hence will try to make sense of it. If the interviewer does not provide additional clarification, they are likely to turn to the wording and context of the ambiguous question to determine its meaning, much as they would be expected to do in any other conversation. Once respondents have assigned a particular meaning to the issue, thus transforming the fictitious issue into a better defined issue that makes sense in the context of the interview, they may have no difficulty reporting a subjectively meaningful opinion. Even if they have not given the particular issue much thought, they may easily identify a broader set of issues to which this particular one apparently belongs, using their general attitude toward the broader set of issues to determine their attitude toward this particular one. Supporting this reasoning, Strack, Schwarz, and Wänke (1991, Experiment 1) observed that German university students favored the introduction of a fictitious "educational contribution" when a preceding question pertained to fellowships for students, but opposed it when it pertained to tuition. Open-ended responses confirmed that they used the content of the preceding question to determine the nature of the fictitious issue.

As the preceding examples illustrate, question comprehension is not primarily an issue of understanding the literal meaning of an utterance. Rather, question comprehension involves extensive inferences about the speaker's intentions to determine the pragmatic meaning of the question. To make these inferences, respondents draw on the nature of preceding questions as well as the response alternatives. Accordingly, survey methodologists' traditional focus on using the "right words" in questionnaire writing needs to be complemented by a consideration of the conversational processes involved in the question-answering process (see Biess, Strack, & Schwarz, 1993; Hilton, 1995; Schwarz, 1994, for a related discussion of psychological experiments).

**Recalling Information and Computing a Judgment** Once respondents determined what the researcher is interested in, they need to recall relevant information from memory. In some cases, respondents may have direct access to a previously formed relevant judgment that they can offer as an answer. In most cases, however, they will not find an appropriate answer readily stored in memory and will need to compute a judgment on the spot. The processes involved in doing so are somewhat different for behavioral questions and attitude questions, and are discussed in the respective sections below.