

Cambridge University Press

0521467330 - Constructing Questions for Interviews and Questionnaires: Theory and Practice in Social Research

William Foddy

Excerpt

[More information](#)

Chapter 1

AN INITIAL STATEMENT OF THE PROBLEM

There is no doubt that the use of verbal data has come to dominate the social sciences. Asking questions is widely accepted as a cost-efficient (and sometimes the only) way, of gathering information about past behaviour and experiences, private actions and motives, and beliefs, values and attitudes (i.e. subjective variables that cannot be measured directly). A review of practices adopted in the 1960s revealed that:

the sociologist today limits himself rather generally to the construction and conduct of questionnaires and interviews. In 1940–41, 25 per cent of the 85 empirical studies depended on interviews and questionnaires for their data; in 1965–66, 48 per cent did. However, if we consider studies based on secondary data that, in turn, derived from interviews, then 64 per cent of the 136 research papers in the latter two years were based on such verbal reports.

Increasingly, these verbal reports are limited to expressions of attitudes, feelings, and opinions rather than to factual accounts of past behaviour and interactions. In 1940–41, 8 of the 22 studies using questionnaires and interviews obtained statements about opinions or feelings, 6 focused on actual behaviour and 8 gathered information about both past behaviour and attitudes. In 1965–66, 49 of 66 studies in which interviews were collected dealt only with opinions and sentiment, 6 with behaviour and 8 with both behaviour and attitudes. It would seem that our colleagues tend to ignore actual behavioural patterns and also fail to come to grips with the fundamental problem of the relation of attitudes and sentiment to behaviour. To an even greater extent, sociology is becoming the study of verbally expressed sentiments and feelings, rather than an analysis of human performance. (Brown and Gilmartin, 1969:288)

2 CONSTRUCTING QUESTIONS

This was the situation in the sixties but there is no reason to believe that the pattern would be any different today — if anything, it is likely to be stronger.

Such an entrenched interest in the use of verbal data would not, in itself, be a bad thing if it could be shown that it always, or even usually, leads to valid conclusions in social research. Unfortunately, it must be admitted that our ability to construct questions which produce data that are reliable and lead to valid conclusions has not been very impressive to date. What is more, recognition of the prevailing situation is far from new. In a report of a survey of expert opinions, Hovde (1936) noted that 74 per cent of the experts who responded mentioned improperly worded questions as a principal defect in commercial social research. The next most frequently mentioned complaint concerned the related issue of faulty interpretations (58 per cent). These figures contrast with 52 per cent mentioning improper statistical methods. The lack of progress since Hovde reported his findings is indicated by Belson's (1986:36) conclusions that the principal causes of error in the gathering of data through survey procedures are:

- (a) respondents' failure to understand questions as intended;
- (b) a lack of effort, or interest, on the part of respondents;
- (c) respondents' unwillingness to admit to certain attitudes or behaviours;
- (d) the failure of respondents' memory or comprehension processes in the stressed conditions of the interview; and,
- (e) interviewer failures of various kinds (e.g. the tendency to change wording, failures in presentation procedures and the adoption of faulty recording procedures).

Examples that illustrate the inadequacy of many of the questions that have been used in social research in the past

It is not difficult to find examples to reinforce the claim that there is a great deal of scope for improving the quality of the data we collect for social research. All ten problems discussed in the next few pages demonstrate this.

Factual questions sometimes elicit invalid answers

Palmer (cited by Deming, 1944) found that, when respondents in a Philadelphia study were re-interviewed eight to ten days after an initial interview, 10 per cent of the reported ages differed by one or more years between the interviews. Likewise, Parry and Crossley (1950) reported that objective checks revealed

Cambridge University Press

0521467330 - Constructing Questions for Interviews and Questionnaires: Theory and Practice in Social Research

William Foddy

Excerpt

[More information](#)

that 5–17 per cent of a random sample of over 900 Denver residents gave incorrect answers to a series of factual questions. The questions included whether or not respondents had registered and voted in various elections, had contributed to the community chest; and possessed library cards and driving licences, as well as details of car ownership.

If questions concerning such simple and apparently objective matters as ‘age’ elicit inaccurate data, one must wonder about the validity problems that might be associated with more threatening, more complex, or less well known issues. A summary of the findings of a number of health surveys published by Marquis (1970) indicates how serious the problem might be. Typically 12–17 per cent of known hospital episodes, 23–26 per cent of recorded visits to physicians, and at least 50 per cent of the chronic and acute conditions listed in medical records were not disclosed by respondents.

The relationship between what respondents say they do and what they actually do is not always very strong

Social scientists have long been tempted to assume that respondents’ behaviour is congruent with their attitudes. Yet the evidence for this link has never been very strong. This issue was first given salience by LaPiere (1934/35) in a paper which has subsequently been discussed by a number of writers (e.g. Deutscher, 1966 and 1973; Phillips, 1971; Schuman and Johnson, 1976). LaPiere spent some weeks during the 1930s travelling around the United States with a Chinese couple. He kept a record of the way in which they were treated at sixty-six hotels and motels in which they had wanted to stay and 184 restaurants and cafes in which they had wanted to eat — only one establishment had refused them service. Six months later, LaPiere wrote to the places in which they had either been given accommodation or meals, asking the proprietors of each establishment if they would accept members of the Chinese race as guests. Fifty per cent replied to his letters and, of these, 90 per cent said ‘No’! This finding focused the spotlight on the apparent fact that respondents do not always do what they say they do.

Findings like those reported by LaPiere have led a number of methodologists (e.g. Cicourel, 1964, 1982; Deutscher, 1966, 1973; Phillips, 1971; Douglas, 1985; Briggs, 1986) to argue that social and social psychological factors which operate in the interview situation invalidate most, if not all, attempts to predict behaviour on the basis of verbally expressed attitudes. The general argument is perhaps extreme; nevertheless, the evidence for a relationship between attitudes and behaviours has always been weak. While this observation might reflect the true nature of the relationship between the concepts, it is also possible that it reflects either a lack of clear conceptualisation of what is being measured and an inadequate theoretical explication of the assumed link between the concepts (see e.g. Weigel and Newman, 1976), or the use of

Cambridge University Press

0521467330 - Constructing Questions for Interviews and Questionnaires: Theory and Practice in Social Research

William Foddy

Excerpt

[More information](#)

4 CONSTRUCTING QUESTIONS

inadequate questions to test hypotheses — this last possibility is discussed in chapter 11.

Respondents' attitudes, beliefs, opinions, habits, interests often seem to be extraordinarily unstable

Twentieth-century social science has predominantly been directed by behaviourist, reductionist and naturalist premises — namely that the objects of inquiry (i.e. 'the stuff out there') has two primary properties: stability and accessibility. And yet available evidence suggests that many sorts of respondents' answers are strikingly variable over time. Converse (1964), for example, reports very low correlations between attitudes expressed by the same respondents over a two-year period. In another study, Bishop *et al.* (1984) found that respondents were less likely to claim that they follow what is going on in public affairs if they had first responded to a set of difficult questions about a congressman's record than if they had to answer the questions about the congressman's record after they had reported their own interest in public affairs. In a third study, Gritching (1986) asked respondents the same question (designed to measure their attitudes toward the setting up of a gambling casino in their community) at the start and at the end of an interview schedule. Gritching reports that 17.6 per cent of the respondents changed their position during the course of the interview.

The results of the studies that have just been reviewed leave us in the position of not knowing whether the observed variation is due to: true variability in the respondents' memory processes (a topic that we will return to in chapter 7), inadequacies or instability in the interpretation of the questions themselves, the impact of cognitions that have been stimulated by earlier questions, the impact of social-psychological variables (e.g. interviewer–respondent status differences — discussed in chapter 9); or the tendency for respondents to spontaneously ask interviewers for their views during interviews (discussed by Oakley, 1981).

Small changes in wording sometimes produce major changes in the distribution of responses

This problem is nicely illustrated by results from Butler and Kitzinger's study of the response distributions for different questions that were formulated by National Opinion Polls to gauge the British people's attitudes toward entering the European Common Market in 1975 (Butler and Kitzinger, 1976:60). On the one hand, the difference between the percentages of 'pro' and 'anti' responses for the question: 'Do you accept the government's recommendation that the

Cambridge University Press

0521467330 - Constructing Questions for Interviews and Questionnaires: Theory and Practice in Social Research

William Foddy

Excerpt

[More information](#)

United Kingdom should *come out of* the Common Market?', was 0.2 per cent in favour of the 'pro' Market position. On the other hand, for the question: 'Do you accept the government's recommendation that the United Kingdom should *stay in* the Common Market?' the difference between the percentages of 'pro' and 'anti' responses was 18.2 per cent in favour of the 'pro' Market position.

A second example that demonstrates the impact that supposedly innocent changes in question wording can have is provided by Bishop *et al.* (1978), who conclude that apparent 'trends' in attitudes over the last thirty years in the United States are likely to be due to changes in the format of the questions that have been used in surveys. Before 1964, the Michigan Survey Research Center asked respondents if they had an opinion about an issue before asking them to rate the issues on a 7-point scale ('No opinion — Agree strongly, Agree but not very strongly, Not sure . . . it depends, Disagree but not very strongly, Disagree strongly, — Don't know'). Between 1964 and 1972, a dichotomised format was used: 'Some people feel X while others feel Y; have you been interested enough to side with X or Y?' After 1972, respondents were told that, 'Some people feel X while others think Y and of course some people have opinions somewhere between', before being asked where they would place themselves on a 7-point numeric scale (with alternative X being paired with 1 and alternative Y being paired with 7). The post-1972 scale does not include either a 'No opinion' or a 'Don't know' response option. Whereas the format used prior to 1964 eliminated one-third of the respondents and the format used from 1964 to 1972 eliminated less than one-sixth of the respondents, the format used after 1972 hardly eliminated any respondents. In other words, after 1972, all respondents were forced to give answers to the questions whether or not they had pre-formed opinions to give, or had been interested in the topic in the past. Bishop *et al.* suggest that differences between the filtering power of each of the formats used (i.e. their ability to eliminate respondents for whom the topic is not relevant) are great enough to account for the apparent changes in public opinion in the United States since the fifties.

A third example that illustrates the effects of what, on the surface, appear to be harmless differences in wording comes from an Australian newspaper (the Melbourne *Age* 6/7/85:3). It was reported that when respondents in a national survey had been asked to rate the performance of the two main federal parliamentary parties on a scale of 'Very good' to 'Very bad', 39 per cent thought the prime minister was doing a good job and 27 per cent thought the same of the leader of the opposition. It was also noted in the same report, however, that two other national polls which had been conducted about the same time had found that 47 per cent approved of the prime minister while 48 per cent approved of his opponent; and 56 per cent saw the incumbent as making the better prime minister, compared with 27 per cent thinking the same for the leader of the opposition. Since the differences between the percentages are too large to be ascribed to sampling differences, it has to be concluded that 'doing a good job',

Cambridge University Press

0521467330 - Constructing Questions for Interviews and Questionnaires: Theory and Practice in Social Research

William Foddy

Excerpt

[More information](#)

6 CONSTRUCTING QUESTIONS

being ‘approved of’ and being seen as ‘making the better prime minister’ must be quite different matters — even though there would have been little reason to suspect this before the polls were conducted.

Last, lest it be thought that only opinion or attitude questions are susceptible to the effects of small differences in wording, it is worth noting that even questions about supposedly simple matters of fact can be vulnerable. Peterson (1984) has reported the results of a study that was designed to compare the non-response rates associated with different ways of asking respondents for their age. Whereas only 3.2 per cent of a random sample of registered voters failed to answer the question ‘What is your age?’, the question ‘How old are you?’, put to a similar sample, produced a 9.7 per cent non-response rate. Again, a researcher would have had no reason to anticipate this outcome before the experiment was carried out.

Respondents commonly misinterpret questions

A number of writers have discussed this problem at length (e.g. Cantril and Fried, 1944; Nuckols, 1953; Cannell, 1977; Belson, 1981, 1986; Hunt *et al.*, 1982). Nuckols (1953) discusses the results of an experiment in which nine questions that had been used by a national polling organisation were re-presented to a sample of respondents who were asked to repeat in their own words their interpretations of them. In all, 17 per cent of the interpretations given were judged to be either partially or wholly incorrect. Belson (1981) investigated respondents’ interpretation of a number of common words including ‘usually’, ‘generally’, ‘people’, ‘children’ and ‘weekday’, and concludes that even these words elicit a wide range of different interpretations. Cannell (1977:44) presents similar data. It would appear that variation in the way respondents interpret everyday words is a common feature of questions used in social research.

Answers to earlier questions can affect respondents’ answers to later questions

Available evidence suggests that ‘contrast’ and ‘consistency’ effects can be generated by prior questions. Rugg and Cantril (1944:28) found that the proportions of respondents agreeing with the idea that Americans should be allowed to enlist in: (a) the French army and (b) the German army, were affected by the order in which the questions concerning the two armies were presented. And Noelle-Neumann (1970) found that, when German respondents were asked to rate various foods in terms of how ‘German’ they were, potatoes

were seen as being particularly ‘German’ by more respondents if that item followed rather than preceded rice.

No doubt it is because of the perceived likelihood of the occurrence of order effects that most methodologists advocate asking general questions about a topic before going on to more specific questions about it. Certainly, answers to prior specific questions often seem to influence answers to later, more general questions, although the nature of this influence is not always the same. Sometimes a consistency effect seems to operate and sometimes a redundancy effect seems to operate. For example, a consistency effect occurs when respondents are asked to indicate how often they pray before being asked to indicate how religious they think they are — that is, respondents who have just reported that they pray a lot are more likely to say that they are very religious (McFarland, 1981). A redundancy effect appears to occur, however, when respondents are asked about how happy they are with their work before being asked about how happy they are in general — that is, respondents seem to exclude consideration of work when answering the second question (see Strack and Martin, 1987; McClendon and O’Brien, 1988; note: we will return to these issues on pages 61–66 of chapter 5).

Changing the order in which response options are presented sometimes affects respondents’ answers

The order in which response options are presented sometimes affects the probabilities of respondents selecting particular options. More specifically, it appears that respondents are more likely to endorse the options that they see first when they are able to read the items for themselves, and more likely to endorse items that they hear last when the items are read out aloud to them (e.g. Krosnick and Alwin, 1987). Notwithstanding these findings, it also appears to be the case that the first response option in a list has a greater impact on evaluative judgements than the last. Thus respondents who are asked to indicate which income category they fall into tend, if the categories are arranged in ascending order, to endorse lower categories than they would if they were arranged in descending order (Locander and Burton, 1976).

Respondents’ answers are sometimes affected by the question format per se

Open ended questions (i.e. questions that allow respondents to supply their own answers) often produce quite different results from closed ended versions of the same questions (i.e. questions that force respondents to select a response

Cambridge University Press

0521467330 - Constructing Questions for Interviews and Questionnaires: Theory and Practice in Social Research

William Foddy

Excerpt

[More information](#)

8 CONSTRUCTING QUESTIONS

from a pre-set list of alternatives). Two issues seem to differentiate the two formats. In the first place, respondents are more likely to endorse a particular option if it has been explicitly listed than they are if they have to spontaneously think of it for themselves. In the second place, it has been found that respondents often give very different types of answers to open questions than they do to congruent closed questions.

The following two examples demonstrate the fact that respondents are more likely to endorse an answer if it has been explicitly listed for them than if it has not. The first is provided by Belson and Duncan (1962:160), who found that a checklist question yielded higher periodical and newspaper readership rates than a comparable open question — for one periodical (the *Radio Times*) the percentage of respondents who checked it on the checklist was over five times the percentage who mentioned it when answering the open version of the question (38% : 7%). The second example comes from Schuman and Presser (1981:86). They report that, whereas 22 per cent of a sample of respondents who had been presented with an open question about what they thought was the most important problem facing the country mentioned the energy shortage, less than 1 per cent of a sample presented with a comparable closed question, which did not specifically include this issue as an option, mentioned it.

That respondents often give quite different kinds of answers to equivalent questions is illustrated by results reported by Schuman and Presser (1981). Some of the respondents, who gave the answer 'the pay' to an open question about what they would most prefer in a job, were subsequently found to have meant 'high pay' while others had meant 'steady pay'.

Respondents often answer questions even when it appears that they know very little about the topic

Respondents frequently answer questions that appear to be marginally relevant to them or about which they have thought little.

It has been found, for instance, that up to 25 per cent of respondents will check substantive options when a 'Don't know' is not offered but check a 'Don't know' option when it is offered (Schuman and Presser, 1981:186). And Ferber (1956) notes that in a random sample of 600 residents of Champaign Urbana, between 14 and 62 per cent of the 50 per cent of respondents who did not know about particular items in an array of topics, still volunteered opinions about them. Similarly, Gallup (1978:1176, cited by Smith, 1984a:221) found that, while 96 per cent of a national sample offered opinions about the importance of a balanced Federal budget, 25 per cent did not know whether the budget was balanced or not, 8 per cent wrongly thought that the budget was balanced, 40 per cent thought that it was unbalanced but did not know by how much, and 25

Cambridge University Press

0521467330 - Constructing Questions for Interviews and Questionnaires: Theory and Practice in Social Research

William Foddy

Excerpt

[More information](#)

per cent thought that it was not balanced but either underestimated or overestimated the amount it was out of balance by more than 15 per cent. In all, Gallup estimates that a mere 3 per cent of the respondents offered an answer that was based on accurate information. Other evidence that relates to respondents' willingness to offer answers that are not well grounded in knowledge comes from studies in which respondents have been asked about either fictitious or extremely obscure topics. Typically, these studies have disclosed that up to 30 per cent of the respondents have been prepared to answer the questions as if they dealt with topics that were real and familiar to them. (See Smith [1984a:223] for a review of the literature.)

The cultural context in which a question is presented often has an impact on the way respondents interpret and answer questions

An example that illustrates the importance of the context in which a question is asked is provided by Briggs (1986), who discusses an ethnic survey that was designed to assess the way a neighbourhood facility could best meet the demands of the residents in the area. He notes that it almost failed because a key question was interpreted quite differently by Navajo respondents, as compared to Zuni, Mexican-American and Anglo-American respondents. The level of demand for services indicated by the Navajo was much lower than that for each of the other groups. While this could have been interpreted as meaning that the Navajo were much less interested than the others in using the services, Briggs realised that the Navajo had not answered the question in the same manner as the other ethnic groups because of their cultural beliefs. Apparently the Navajo deem it highly inappropriate to speculate about the beliefs held by others. In their culture, such talk is seen as a usurpation of others' decision-making powers.

Speculating on the preferences of one's spouse and children would accordingly be deemed extremely rude. Rather than doing so, Navajo respondents would estimate which services they themselves would be likely to use. The use of a probe to obtain data on the family members generally yielded statements such as 'no, I don't think so . . .'. (Briggs, 1986:97)

Critiques of past practices

Perhaps because of the sorts of problems discussed above, several trenchant critiques of the use of questionnaires and interviews in social research have been published over the last two decades (see, e.g., Deutscher, 1966, 1973; Phillips,

Cambridge University Press

0521467330 - Constructing Questions for Interviews and Questionnaires: Theory and Practice in Social Research

William Foddy

Excerpt

[More information](#)

10 CONSTRUCTING QUESTIONS

1971; Cicourel, 1964, 1982; Douglas, 1985; Mishler, 1986; Briggs, 1986; Pawson, 1989). Cicourel sums up the situation in the following way:

Despite the fact that virtually all social science data are derived from some kind of discourse or textual materials, sociologists have devoted little time to establishing explicit theoretical foundations for the use of such instruments as interviews and surveys. A key problem has been the lack of clear theoretical concepts about the interpretation of interview and survey question and answer frames. We lack a theory of comprehension and communication that can provide a foundation for the way question-answer systems function . . . (Cicourel, 1982:Abstract)

Yet for all the criticism, reliance on verbal data does not appear to be in any danger of waning. Indeed many critics (e.g. Phillips, 1971; Cicourel, 1982; Briggs, 1986 and Pawson, 1989) have implicitly recognised this and have directed their analyses toward improving rather than demolishing current methodological practices. Cicourel acknowledges the value of interviews for many 'theoretical purposes' and takes for granted that a number of old issues have been generally resolved. He assumes, for example, that it is generally accepted that qualitative interviews should normally precede the construction of formalised questions. He also assumes that it is necessary to pre-test the resulting questions to ensure both that respondents understand them and that answer categories reflect respondents' thinking. Briggs continues in the same vein by suggesting several procedures which, if followed, might help to sensitise the researcher to the communicative norms that prevail in the respondents' community. He advocates, for instance, that the researcher undertake preliminary pilot work to discover who talks to whom; who listens when individuals talk and when they remain silent; the ways in which individuals communicate with one another; how questions are asked; and what sort of questions are seen as legitimate. He suggests, too, employing other procedures — for example, video tapes — to help discover how, and what, individuals communicate; to allow the researcher to look at the same interactions several times to identify underlying assumptions or perspectives adopted by interactants, and to allow respondents to reflect and comment upon the researcher's interpretations of their interactions. And Phillips stresses both that social psychological factors (e.g. interviewer–respondent social status differences) have to be taken into account and that interview behaviour is a form of social interaction and should be analysed accordingly.

Thus, Cicourel, Briggs and Phillips all imply that question–answer situations are more complex than have generally been appreciated. Their arguments suggest that questions and answers should be understood in relation not only to one another but also to the social contexts in which they operate. The important implication of these arguments is that it is not enough to arbitrarily limit attention to aspects of questions themselves; to the way properties of questions