Chapter 9

Sir Claus Moser and Graham Kalton

QUESTIONNAIRES

From *Survey Methods in Social Investigation* (2nd edition), Aldershot: Gower (1971).

. . .

THE FIRST STEP IN designing a questionnaire is to define the problem to be tackled by the survey and hence to decide on what questions to ask. The temptation is always to cover too much, to ask everything that might turn out to be interesting. This must be resisted. Lengthy, rambling questionnaires are as demoralizing for the interviewer as for the respondent, and the questionnaire should be no longer than is absolutely necessary for the purpose. Certain questions will, so to speak, include themselves, but a problem of choice inevitably arises with marginal ones. Let us consider a hypothetical survey to ascertain what daily newspapers different kinds of people read. A number of newspaper questions, together with those asking for necessary personal data, automatically suggest themselves. Then, as the discussion of the planning of the survey warms up, many extensions of interest occur to those taking part. Would it be useful to include reading of periodicals and books? Would the main results be more meaningful if they could be viewed against the background of the respondents' leisure habits as a whole? Would it be wise to find out something about how much money and time different people have available for newspaper buying and reading? Should one ask a question or two about the use of libraries? Should one go beyond the facts of reading and ask people's opinions on individual newspapers?

And so it goes on, with the questionnaire growing from a short list of questions to a document many pages long. . . . It is obvious that the survey planner must rigorously examine every question, and exclude any that are not strictly relevant to

the survey's objectives. In this, the pilot survey is his most helpful tool. Here all the marginal questions can be tested out and dummy tabulations made from the results. Questions likely to prove of small importance in the final analysis can be spotted, as can those which turn out to be not worth asking unless a host of others is also included.

In settling the scope of a questionnaire, one other criterion should be applied, namely that the questions should be practicable. This merits emphasis, even though no amount of textbook admonition can take the place of common sense. It is no good asking a person's opinion about something he does not understand; about events too long ago for him to remember accurately; about matters which, although they concern him, he is unlikely to have accurate information on or that are so personal or emotional that valid answers cannot be expected by formal direct questioning.

Question content

In considering any question, then, it is wise to ask oneself whether respondents are likely to possess the knowledge, or have access to the information, necessary for giving a correct answer. It is unsafe to *assume* that respondents will voluntarily admit ignorance. On the contrary it has often been shown that they will give some kind of answer to most questions, even if they are ill-informed and know it. Similarly, they will express opinions on matters they have given little thought to or which they barely understand. . . .

The surveyor should aim to ask questions only from those likely to be able to answer them accurately; to ask about past events only if he can reasonably expect people to remember them accurately (perhaps with the help of recall methods); and to ask their opinions only if he can be reasonably sure that they understand what is involved and are able to give meaningful answers. It is always well to remember that most survey questions are addressed to a variety of people very differently qualified to answer them. . . .

Most survey questions are concerned with either facts or opinions. There are also questions dealing with motivation ('Why did you go to the cinema last night?'), and knowledge questions ('What do the initials NATO stand for?'), but the main points of methodology will emerge if we consider factual and opinion questions. . . .

The chief difficulties with factual questions are to ensure that interviewers understand, and manage to convey to the respondents, precisely what facts are wanted. Some of the definitions may be tricky but, in most cases, the chances of either interviewer or respondent misunderstanding the question, not understanding it at all, or the latter being influenced in his answer by the words chosen are much slighter than with opinion questions.

With opinion questions the problems are much more fundamental. Though we would not venture into the psychologist's territory and discuss concepts of opinion and attitude in any detail, some attempt must be made to analyse why the study of opinions is basically so much more troublesome than that of facts. Why would one be more confident with a question asking a respondent whether he owns a

wrist-watch than with one asking whether he is in favour of capital punishment? There are several related reasons:

(*a*) A respondent either does or does not possess a watch, and one may reasonably assume that he knows whether he does or not. All the surveyor has to do is to make clear to the respondent what he wants to know, and to be sure he understands the respondent's answer. It may be that the respondent wishes not to give the correct answer, but at least he knows what it is. With the opinion question it is not so simple. The respondent's attitude to capital punishment may be largely latent, and he may never have given the matter any conscious thought until he was confronted by the question. The first problem with opinion questions thus arises from the uncertainty whether the respondent, in any meaningful sense, 'knows' the correct answers. To say whether he possesses a watch or not needs no 'thinking' on the respondent's part; to give a genuine opinion on capital punishment may require thought and 'self-analysis'.

(*b*) A person's opinion on virtually any issue is many-sided. On capital punishment there are moral, medical, legal and other aspects: it is possible to be against it on moral grounds, in favour on legal ones. A person may be against it in all but certain circumstances, or against it whatever the situation. He may be in favour of abolishing it experimentally, or as an irrevocable step whatever the consequences. In short, there probably is *no one correct answer* to the survey question as there is to that on watch ownership. The answer the respondent actually gives will depend on the aspect of the issue that is uppermost in his mind – quite possibly because the wording of the question, or the context created by previous ones, has put it there.

(*c*) Closely related to this is the problem of intensity. On any given subject some people feel strongly, some are indifferent, some have settled and consistent views, others are highly changeable in their attitude. In any attempt to get more than snap answers, the problem of assessing the intensity of opinion and attitude must be faced.

(*d*) Finally, it must be repeated that answers to opinion questions are more sensitive to changes in wording, emphasis, sequence and so on than are those to factual questions. . . . This established sensitivity of opinion questions does not imply instability of opinion among respondents. Rather it is a reflection of the point made in (*b*) above. Opinion is many-sided, and questions asked in different ways will seem to 'get at' different aspects of the opinion: if they result in different answers, it is largely because respondents are in effect answering different questions.

There is a secondary difficulty here. With factual questions, it is often feasible to compare the merits of different forms of the same question by checking the answers against known data. With opinion questions this is impossible, although checks on validity can and should be made; where, for instance, opinions are closely related to measurable behaviour, a check on behaviour can be used to test the validity of an expressed opinion. . . .

Question wording

The literature on the wording of questions is bewildering. Numerous papers have appeared showing the relative advantages of various specific questions, the danger of using a certain word or phrase, the sensitivity of answers to changes in wording and presentation: but it is exceedingly difficult to build out of them any general principles. We shall confine ourselves to some aspects of wording which are of general importance in social research surveys.

(a) *Questions that are insufficiently specific.* A common error is to ask a general question when an answer on a specific issue is wanted. If one is interested specifically in a canteen's meal prices and the quality of its service, the question 'Are you satisfied or dissatisfied with your canteen?' is unsatisfactory, since it fails to provide the respondent with the necessary frames of reference. As there are two distinct frames of reference of interest here, two questions are needed, perhaps 'Are you satisfied or dissatisfied with the prices of meals in your canteen?' and 'Are you satisfied or dissatisfied with the service in your canteen?' Although these two questions now cover the topics required in a seemingly straightforward way, they still need to be pre-tested to check on their suitability for the particular situation. It may, for instance, be the case that the canteen serves special meals once a week at a higher cost and that, although generally satisfied with the canteen's prices, a respondent objects to the cost of the special meals; or he may be dissatisfied only with one particular aspect of the service. In cases like these he would have difficulty answering the questions. Such problems are brought to light by pre-testing and pilot work, the importance of which for question wording cannot be overrated. . . .

(b) *Simple language.* In choosing the language for a questionnaire the population being studied should be kept in mind. The aim in question wording is to communicate with respondents as nearly as possible in their own language. A survey of the members of a particular profession, for instance, can usefully employ the profession's common technical terms; not only are such terms part of the informants' common language, but they also normally have a single precise meaning, unlike everyday terms, which particularly to professionals are often vague and ambiguous.

Technical terms and jargon are, however, obviously to be avoided in surveys of the general population. We would not ask a respondent whether his household is run on matriarchal lines, what he thinks about bilateral trading, amortization of the National Debt, and fiscal policy.

Much less easy to recognize and reject are words which, though everyday usage to the university-trained survey expert, are far from common in ordinary conversation. Words like hypothetical, irrespective, aggravate, deprecate, and hundreds more are in this category. . . .

With surveys of the general population, the first principles in wording are that questions should use the simplest words that will convey the exact meaning, and that the phrasing also should be as simple and informal as possible. It is more natural to ask: 'Do you think . . . ?' than: 'Is it your opinion . . . ?'; 'What is your attitude to . . . ?' than: 'What is your attitude with regard to . . . ?' In fact the more questions sound like ordinary conversation the smoother the interview will be. Of course, this

should not be overdone. Bad grammar may be more common than good, but one would not advocate its deliberate use in survey questions. Nor are slang expressions advisable; as with technical jargon, not everyone uses the same expressions. It is not indeed enough to know that a word or phrase is commonly used; one must equally be sure that it is used in the same sense by all groups of respondents. Even words like 'dinner' and 'tea' have different meanings in different parts of the country. A simple case is the word 'book', which in some parts of the population is taken to include magazines. Hence the phrasing of the following question in a readership survey by Stuart (1952): 'During the past week roughly how many hours would you say you had spent reading *books* – I mean books not magazines or papers?' . . .

There is a temptation to ask complex questions when the subject matter is inherently complicated, involving a variety of different facets. This, for example, would be the case in a housing survey in which one wanted to discover how many households comprised three-generation families, that is grandparents, parents and children. Once the term 'three-generation family' has been precisely defined (how about widowed grandparents, unmarried mothers, divorced or separated parents?), one might with ingenuity design a single question to obtain the information, but many respondents would certainly fail to understand it. Rather than rely on a single complex question, a series of simple questions should be asked, the number of such questions depending on the degree of simplicity required. Household composition is generally a complex subject and one for which several descriptive indices are required; the information is usually best obtained by using a 'household box' on the questionnaire in which the household members are listed together with their relevant characteristics, e.g. age, sex, marital status, working status and educational level. . . . From these basic data the surveyor can determine for himself all the indices he requires for his analysis. . . .

(c) Ambiguity. Ambiguous questions are to be avoided at all costs. If an ambiguous word creeps in, different people will understand the question differently and will in effect be answering different questions. The following example is taken from a university research survey: 'Is your work made more difficult because you are expecting a baby?' The question was asked of all women in the survey, irrespective of whether they were expecting a baby or not. What, then, did a 'No' answer mean? Depending on the respondent, it might have meant 'No, I'm not expecting a baby' or 'No, my work is not made more difficult by the fact that I'm expecting a baby'.

Ambiguity also arises with double barrelled questions, such as the following question on public transport: 'Do you like travelling on trains and buses?' Respondents liking one and disliking the other would be in a dilemma in answering this question. Clearly it needs to be divided into two questions, each concerned with a single idea, in this case with a single mode of transport.

(d) Vague words. Vague questions encourage vague answers. If people are asked whether they go to the cinema regularly or occasionally, the meaning of their answers will be vague. (This common choice of alternatives is strictly illogical; the word 'occasional' refers to frequency, the word 'regular' does not. However, this may be a case where logic can give way to common usage.) The meaning can

easily be made more precise, as in the following question from the 1968 National Readership Survey: 'How often these days do you go to the cinema? Would it be nearer to – twice a week or more often; once a week; once a fortnight; once a month; three or four times a year; less often; or do you never go these days?'

Vague words and phrases like 'kind of', 'fairly', 'generally', 'often', 'many', 'much the same', 'on the whole' should be avoided, unless one is only seeking vague answers. If one asks 'What kind of house do you have?' without specifying a frame of reference, some people will answer that it is semi-detached, others that it is suburban, others that it is very pleasant, and so on.

A similar type of vagueness occurs in 'Why' questions. In answering the question: 'Why did you go to the cinema last night?' some respondents will say that they wanted to see that particular film, some that they did not want to stay at home, others that 'the wife suggested it', or that they hadn't been since last week. The word 'Why' in this question – as the phrase 'kind of' in the previous one – can mean so many different things that its use would produce a useless mixture of answers. Lazarsfeld (1935) discusses the problems of the 'Why' question.

(e) Leading questions. A leading question is one which, by its content, structure or wording, leads the respondent in the direction of a certain answer. The question form: 'You don't think . . . do you?' as obviously leads to a negative answer as the form: 'Should not something be done about . . . ?' leads to a positive one.

Equally, a question which suggests only some of the possible answers may lead in their direction. Take the question: 'Do you read any weekly newspapers, such as the *New Statesman* or *Punch*?' Respondents, especially if they are not sure of their correct or complete reply, may seek refuge in the answers named; either all or none of the alternatives should be stated.

There are numerous words that have been shown on occasion to have a 'leading' influence in survey questions (see Payne 1951, and Cantril 1944). The word 'involved' in a question like: 'Do you think that the Government should get involved in . . . ?' may have a sufficiently sinister ring to lead people in the negative direction. Similarly, the wording: 'Do you agree that the Government is right in staying out of . . . ?' invites a 'Yes' answer. The 'leading' nature of these examples is obvious, but more subtle leads can often creep unnoticed into survey questions . . .

(f) Presuming questions. Questions should not, generally speaking, presume anything about the respondent. They should not imply that he necessarily possesses any knowledge or an opinion on the survey subject, or that he engages in the activity about which he is being asked. Questions like: 'How many cigarettes a day do you smoke?' or 'How did you vote in the last General Election?' are best asked only after a 'filter' question has revealed that the respondent does smoke cigarettes and did vote in the last election.

On occasion, however, one might deliberately depart from this procedure. Kinsey and others (1948) did not first ask respondents *whether* they had engaged in certain sexual practices, but went straight into questions about frequency and detail. Respondents were thus spared the embarrassment of admitting the experiences directly and were made to feel that these represented perfectly usual behaviour: thus they found themselves able to talk freely and give detailed answers. The case for

such an approach is obvious, but one cannot ignore the possibility that it may discourage 'I never do' answers and thus cause an upward bias in the results.

(g) Hypothetical questions. Questions of the 'Would you like to live in a flat?' type are of very limited value. Most people would like to try anything once, and an affirmative answer would have little value as a prediction of behaviour. It is another matter if one has first made sure that the person has experience of both flat and house dwelling. Equally, answers to the 'What would you do if . . . ?' kind of question, although perhaps a good reflection of wishful thinking or of what people feel to be right and proper, are unsafe pointers to future behaviour.

Yet prediction of future behaviour on the basis of survey questions plays, and must be expected to play, a central role in survey applications. Market researchers would like – and try – to predict how people will react to a proposed change in the price of a product, to an alteration to its quality or packaging; how many people are likely to buy cars, radios or television sets in a given period, and so on. They may rely on straight questions (a Gallup Poll question in 1950 was: 'Supposing the price of (a certain newspaper) went up from 1d. to 1½d. would you change over to another paper where the price hadn't gone up?') but the answers are recognized to be imperfect guides to future behaviour. People are not good at predicting their behaviour in a hypothetical situation and the prediction has somehow to be taken out of their hands and made by the researcher himself – naturally on the basis of the information he has obtained.

Another kind of hypothetical question is 'Would you like a more frequent bus service?' or 'Would you like an increase in wages?' Such questions are unlikely to be of any value because the respondent is being asked if he would like something for nothing. It is hard to see how he could possibly say 'No'. If he did, it could only be because he has taken into account some hidden factors of his own, or because he has failed to understand the question.

(h) Personalized questions. It is often necessary to decide whether a question should be asked in a personalized form or not. This is well illustrated by the following questions which appeared, one after the other, in a schedule dealing with health matters (see David, 1952): 'Do you think it is a good idea to have everyone's chest regularly checked by X-ray?' and 'Have you ever had yours checked?' Some 96 per cent of the respondents answered 'Yes' to the first question, but only 54 per cent to the second. As the author suggested, the opinion given in answer to the first question 'is more a pious hope for some vague corporate decision than a considered aim involving personal action'.

(i) Embarrassing questions. Subjects which people do not like to discuss in public present a problem to the questionnaire designer. Respondents are often embarrassed to discuss private matters, to give low-prestige answers, and to admit to socially unacceptable behaviour and attitudes. If, for instance, questions on sexual behaviour, frequency of taking a bath, cheating in examinations or attitudes to Communism were asked in the usual way, many respondents would probably refuse to reply and others would distort their answers. There are several ways of attempting to deal with this problem.

One method of reducing the threatening nature of a question is to express it in the third person; instead of asking the respondent for his views, he is asked about the views of others. An example from market research of an indirect question of this sort is given by Smith (1954): 'Some women who use this cleanser find a lot of faults with it. I wonder if you can guess what they are objecting to'. The purpose of this wording was to make the housewives feel free to criticize the product. The aim of such questions is to obtain the respondent's own views, but he may of course answer the question asked, and give what he believes to be the views of others. For this reason it is often advisable to follow the indirect question by a direct one asking the respondent whether he holds the views he has described.

There are several other indirect methods which can be useful in dealing with embarrassing topics. The respondent can for instance be shown a drawing of two persons in a certain setting, with 'balloons' containing speech coming from their mouths, as in comic strips and cartoons. One person's balloon is left empty and the respondent is asked to put himself in the position of that person and to fill in the missing words. Another method is that of sentence completion; the respondent is given the start of a sentence and is asked to complete it, usually under time pressure to ensure spontaneity. Oppenheim (1966) describes the use of the following two examples of sentence completion in a study among psychiatric nurses in a mental hospital:

> 'I wish that doctors . . .'
> 'Patients who are incontinent . . .'

The different ways in which a group of student nurses and a group of nurses with twelve or more years of experience completed these sentences showed the difference of attitude and approach of the two groups. . . .

Belson (1968) describes a study of a randomly derived sample of London teenage boys on the sensitive subject of stealing. A variety of procedures were employed in this study to make it easier for the boys to admit that they had stolen things. On arrival at the interviewing centre a boy chose a false name and, in order to preserve his anonymity, he was introduced under his false name to the interviewer, who knew him only by that name. After an extended initial phase, the interview proceeded to the card-sorting technique by which the information on stealing was to be obtained. The interviewer and the boy sat on either side of a table, with a screen in between so that they could not see each other. Through a slot in the screen the interviewer passed to the boy a card on which one type of stealing (e.g. 'I have stolen cigarettes') was recorded. The boy was asked to put the card in a box labelled 'Yes' if he had ever done what was recorded on it, and in a box labelled 'Never' if not. This was repeated for 44 kinds of theft. At the end of this sorting stage, the interviewer went through a procedure which tried to reduce the force of a boy's resistances, and to strengthen his feeling of willingness, to admitting thefts. Then the boy was asked to re-sort all the cards he had put in the 'Never' box. Finally he was asked for further details on each type of theft he had admitted. This detailed procedure elicited reports of many types of theft from many boys with, for example, 69 per cent of boys admitting 'I have stolen something from a shop' and 58 per cent 'I have stolen money' at least once in their lives. . . .

(j) Questions on periodical behaviour. An interesting choice arises in studying the frequency of periodical behaviour. The main choice of questions can be illustrated with reference to cinema-going:

(i) 'How often have you been to the cinema during the last fortnight (or any other period chosen)?'
(ii) 'How often do you go to the cinema on the average?'
(iii) 'When did you last go to the cinema?'

The first question covers a number of different possibilities corresponding to the period chosen, and answers will depend on the type of activity and on the extent to which one is willing to rely on the respondent's memory (see *(k)* below). In any case, the three question types might produce different results, and there is little evidence on which to choose between them. At first sight, (i) seems to be most specific, but many people's answers might simply be an estimate of their average cinema-going rather than the actual figure; i.e. if they normally go twice a fortnight, they may give this as an answer, although they went only once in the last fortnight. As a case in point, Belson (1964a) reports that an intensive interview follow-up enquiry of respondents to the National Readership Survey suggested that people frequently answered in terms of what publications they *usually* looked at, rather than what they had *actually* looked at, which was what was required. Of course the two answers will often be the same, and it is only when a difference arises that an answer in the wrong terms produces error.

Many survey questions involve this type of choice, e.g. questions on newspaper reading, radio listening, television watching, and consumer purchases. It is a matter deserving further research.

(k) Questions involving memory. Most factual questions to some extent involve the respondent in recalling information. His degree of success in doing this accurately is thus a basic determinant of the quality of his response. With certain questions, such as 'Are you married, single or widowed?', there is no such problem, but with a large range of survey questions recalling information does present a problem, the severity of which depends on what is to be recalled. Two factors of primary importance in memory are the length of time since the event took place and the event's importance to the respondent; events the respondent considers insignificant are likely to be forgotten almost immediately and even the recollection of significant events decreases as time elapses. Moreover, for events not forgotten in their entirety, memory acts selectively, retaining some aspects and losing others, thus producing distorted images. For questions dealing with the past, serious attention must therefore be given to the respondents' abilities to recall the required information accurately, and to ways by which they can be helped to do so.

As an example . . . a memory problem arises with questions asking respondents to provide a list, as would be the case for instance if they were asked which television programmes they had viewed yesterday, or which newspapers they had read or looked at in the preceding seven days; without help many respondents would be unable to give a complete list. A sensible way to aid recall in this case is to provide the respondent with a list of all television programmes transmitted yesterday (or a

list of all newspapers), from which he can pick out the ones he had seen (or read, or looked at). In the National Readership Surveys, for example, respondents are asked about their readership of each publication from a complete list of every publication with which the surveys are concerned. With the interviewer they go through booklets containing the title blocks of the publications, and are asked about each one in turn. The use of the title blocks in these recall-aid booklets is an example of another useful device, visual aids, to assist recall.

With questions like the readership one, there are two types of memory error. The first is the 'recall loss', occurring when the respondent fails to report an activity in the recall period because he has forgotten about it, and this loss is likely to be more serious the longer the period. The second occurs when he reports an activity in the recall period when it actually took place outside that period: the tendency to report as occurring in the current period events which in fact occurred earlier has been termed the 'telescoping effect'. A greater telescoping effect for shorter recall periods has been suggested as part of the explanation for the commonly found effect of relatively greater reporting rates for short recall periods. . . .

With serious memory errors having been demonstrated in many studies, it is natural to look for a procedure which does not rely heavily on an informant's ability to recall information. One obvious possibility is to persuade informants to keep diaries of the events of interest, as is done in the Family Expenditure and National Food Surveys. Diaries, however, have their own limitations. First, the amount of work asked of the respondent is much greater with the diary method, and this may make it difficult to gain the co-operation of the selected sample – the refusal rate may be high. Secondly, the diary method is likely to be more expensive, for interviewers will probably need to contact informants at least twice. One visit is needed to gain the informant's co-operation and to explain the recording procedure, and another is needed to collect the completed diaries. During the recording period other visits may be made to ensure that the instructions have been understood, to check that the data are being correctly recorded, and to maintain morale. The last visit serves not only for the collection of the diary, but also as an opportunity for the interviewer to edit the diary with the respondent; were it not for this editing, the last visit could perhaps be dispensed with, for the diaries could be returned by post.

Even with careful editing, however, the standard of informants' recording cannot be expected to reach that achieved by well-trained interviewers. Surveys of the general population contain people from a wide range of educational levels and with varying amounts of form-completing experience; it can be anticipated that some of them will fail to understand from one interview exactly what they are to do. In addition, others may lack the motivation to complete the diaries as accurately as is required. One particular way in which informants may deviate from instructions is by failing to record the events while they are fresh in their memories; the main strength of the diary approach is the avoidance of reliance on memory, but, if the informant does not keep the diary up-to-date, at least part of that strength is lost. Another source of error is that, although instructed not to change their habits as the result of their recording, some informants will do so; in consumer expenditure surveys, for instance, housewives keeping log-books of their purchases may become more aware of their shopping habits, and this may for example persuade them of the advantages of buying larger items and of shopping in supermarkets. . . .

These limitations of the diary method must be balanced against the memory errors involved in the recall method. The choice between the methods depends on the subject matter of the survey and, in particular, on the ability of respondents to recall accurately the necessary details of the information required. In situations where, even with assistance from the interviewer, informants are unable to recollect details accurately, the recall method is inappropriate and the diary method may be the only possible approach.

Open and pre-coded questions

The relative merits of open and pre-coded questions have been the subject of a good deal of research and debate. In an open question the respondent is given freedom to decide the aspect, form, detail and length of his answer, and it is the interviewer's job to record as much of it as she can. In the case of pre-coded questions, either the respondent is given a limited number of answers from which to choose or the question is asked as an open question and the interviewer allocates the answer to the appropriate code category. . . .

The essential difference thus lies in the stage at which the information is coded, whether in the office, by the respondent or by the interviewer. If the researcher wants a very detailed answer, or wishes to find out what aspects of an issue are uppermost in the respondent's mind, [an open question] is to be preferred. Even if it has to be summarized subsequently, all the detail is there, not merely a number representing the nearest code answer. Any summarizing or coding can be carried out uniformly in the office, uninfluenced by the circumstances of the interview or the reaction of the respondent to the interviewer. But, of course, open questions have their problems. The detail obtained is partly a reflection of the respondent's loquacity, so that different amounts (as well as different items) of information will be available for different people. A second difficulty lies in the task of compressing a written, qualitative answer into code categories. Again, although the remoteness of the office from the interview situation ensures some gain in coding objectivity, it also has drawbacks. Just as questions can sound different if asked by different people, so the meaning of an answer is communicated partly by the way it is given, and this will not be reflected in the written record. Finally, there is the difficulty of getting a verbatim report of what is said. All interviewers probably exercise some selection in recording answers and, to the extent that this happens, bias may creep in.

Pre-coded questions may offer two or more alternative answers (referred to respectively as dichotomous and multiple-choice – or 'cafeteria' – questions) and their advantages are evident. To combine the recording and coding of answers in one operation simplifies the whole procedure; and, in a very real sense, the interviewer is the person best placed to arrive at an accurate coding, since she hears the answers in full and thus has more data to work on than the office coder. On the other hand, once she has ringed a code there is little hope of detecting errors of recording or judgement. Also, she is working under pressure and may be unable to give sufficient time and attention to the needs of a complex coding operation.

If the range of answers to a question is limited and well established, pre-coding is generally to be preferred. Most factual questions – with regular exceptions like

questions on occupation – belong to this category. If, however, one cannot reasonably determine in advance what the main categories will be, it is best to begin with open questions, progressing to pre-coded ones as the range and distribution of answers become clear. This is why open questions play such a valuable role in pilot surveys.

The alternatives offered in pre-coded questions must above all be exhaustive and mutually exclusive. (The code 'Other, specify . . .' is usually added for rare or unthought-of answers.) In such questions all the possible answers must be given. The following question occurred in an opinion survey: 'What happens to the copy of the . . . (newspaper); for instance, does anyone take it to work?'

Stays in house	1
Regularly taken to work, left there	2
Occasionally taken to work, left there	3
Taken to work, brought home	4

It is likely that the form of the question disfavoured the first code answer. If any of the answers are to be suggested, *all* should be. A respondent who has never considered the subject of the question carefully may seize upon any lead given by the mention of a possible answer.

A risk with pre-coded questions is that answers may be forced into a category to which they do not properly belong. Take the hypothetical question: 'Do you think the present Government is doing a good or bad job?' Many people will have clear views and will unhesitatingly say 'Good' or 'Bad'. But what of those who are inclined to say 'Good, but . . .' or 'Bad, except that . . .'? The coding demands a decision one way or the other and may result in qualified responses being forced into categories to which they do not genuinely belong. To try to avoid this, survey designers leave space for qualifications or allow in the codes for finer shades of opinion. Up to a point, a greater number of codes has the added advantage that more information is collected. But there is a limit: if too many codes are used, respondents will be unable to make a rational choice between several of the alternatives and, faced with so many codes, they may have difficulty in making a choice at all.

Besides fixing the number of codes to be used, with opinion questions the survey designer has also to decide whether or not to code for a neutral position, in other words he must decide whether he wants to force respondents to come down on one side or other of the fence. If he does provide a neutral code, he may well find that many people take up that option. The following question was included in a schedule on saving habits: 'During the coming year do you think things will get much better or worse for people in your position or do you think there is not likely to be much change?' The last phrase offered a neutral escape, and 44 per cent of the respondents chose it. These answers may of course express genuine opinions, but there is clearly a risk in suggesting a non-committal answer to the respondent. . . .

Question order

In putting the individual questions together to form the questionnaire, the order of questions needs to be planned. The order may affect the refusal rate and there is plenty of evidence that it may also influence the answers obtained (e.g. Mosteller and others 1949, Cantril 1944, Whitfield 1950), especially so when one is concerned with opinions that are unstable or marginal.

At the start of the interview the respondent is unsure of himself and so the opening questions should be ones to put him at ease and build up rapport between him and the interviewer. They should be interesting questions which he will have no difficulty in answering, and they should not be on sensitive topics, for otherwise he may refuse to continue with the interview. The questions should then proceed in a logical manner, moving from topic to topic in a way that indicates to the respondent the relationship between the questions; where an obvious break in subject matter occurs it is usually advisable to give a sentence or two explaining the break and the relevance of the new set of questions. Since questions on highly sensitive topics may lead to the respondent refusing to continue with the interview, they may be best left until last; then, if a refusal is met, relatively little information is lost.

When determining the order of questions within a topic (and also, for that matter, between topics) the conditioning effect of earlier questions should be considered. It is no good asking: 'Can you name any washing powder?' if a previous question has mentioned 'Tide' or 'Dreft'; in other words knowledge questions must not be preceded by others giving relevant information. Even though interest may centre on specific issues, it can be a good idea to start with broad questions about the subject and then to narrow down to the specific issues, using what is known as a *funnel sequence* of questions (Kahn and Cannell 1957). Thus a general open question on the achievements of the present Government may be the beginning of a sequence leading to specific questions on the Government's actions in the field of labour relations; a mention of labour relations in reply to the first question suggests that the respondent attaches some importance to the subject. On the other hand, if one is interested in the broader question and one thinks the respondents do not hold considered opinions about it, an inverted funnel sequence may be useful. In this case, the early questions ask about the range of issues involved and, in answering them, the respondent is led towards forming a considered opinion on the broader question.

A fairly common situation is one in which the respondent is taken through a list of items by the interviewer, who asks the same initial question about each item in turn. If the respondent answers this question in a certain way the interviewer asks supplementary questions: otherwise she proceeds to the next item. Respondents soon learn in this situation that they can complete the interview more rapidly by avoiding the replies leading to supplementary questions, and this may tempt some to falsify their replies. This risk is easily avoided, however, by asking the supplementary questions only after answers to the initial question have been obtained for all the items on the list.

Another problem with long lists of items is that of respondent fatigue: towards the end of a list of, say, 90 items (about the number of publications in the National

Readership Surveys) the respondent can be expected to experience fatigue, which may result in him failing to recall the later items and hence answering the questions about them negatively. In the National Readership Surveys, for instance, it has been found that the readership level for the group of weeklies when they appeared last in the presentation order (after the groups of dailies, Sundays and monthlies) was only about three-quarters of that when they appeared first (Belson 1964); fatigue probably provides at least a partial explanation of this finding. In these surveys, to avoid bias arising from the order of presentation, the order of the four groups is varied by a rotation scheme throughout the sample; in addition, for one half of the sample the publications within a group are presented in one order and in the other half in the reverse order. This procedure may mean that somewhat better comparisons can be made between the readership levels of different publications, because they have on average about the same presentation position (although account must also be taken of the variation in the 'rotation effect' for the different publications), but it does not make the absolute readership levels for all publications more accurate.

Concluding remarks

We have not attempted to deal comprehensively with the subject of question wording. The points selected for discussion have been those thought to be of most interest to the student or researcher embarking on a survey. To the problem of questionnaire design in general there is no easy solution. Even if one follows all the accepted principles, there usually remains a choice of several question forms, each of which seems satisfactory. Every surveyor tries to phrase his questions in simple, everyday language, to avoid vagueness and ambiguity and to use neutral wording. His difficulty lies in judging whether, with any particular question, he has succeeded in these aims. He may appreciate perfectly that leading questions are to be avoided, but how can he know for sure which words will be 'leading' with the particular question, survey and population that confront him, perhaps for the first time?

The answer to this question lies in detailed pre-tests and pilot studies: more than anything else, they are the essence of a good questionnaire. However experienced the questionnaire designer, any attempt to shortcut these preparatory stages will seriously jeopardize the quality of the questionnaire; past experience is a considerable asset, but in a fresh survey there are always new aspects which may perhaps not be immediately recognized, but which exist and must be investigated through pre-tests and pilot studies. . . .

Question designing remains a matter of common sense and experience and of avoiding known pitfalls. It is not as yet, if indeed it ever can be, a matter of applying theoretical rules. Alternative versions of questions must be rigorously tested in pre-tests and the pilot survey, for in the absence of hard and fast rules, tests of practicability must play a crucial role in questionnaire construction.

References

Belson, W.A. (1964) 'Readership in Britain'. *Business Review* (Australia), **6**, 416–420.

—— (1968) 'The extent of stealing by London boys and some of its origins'. *Advancement of Science*, **25**, 171–184.

Cantril, H., ed. (1944) *Gauging Public Opinion*. Princeton University Press, Princeton.

David, S.T. (1952) 'Public opinion concerning tuberculosis'. *Tubercle (Journal of the British Tuberculosis Association)*. **33**, 78–90.

Kahn, R.L. and Cannell, C.F. (1957) *The Dynamics of Interviewing; Theory, Technique, and Cases*. Wiley, New York.

Kinsey, A.C., Pomeroy, W.B. and Martin, C.E. (1948) *Sexual Behavior in the Human Male*. Saunders, Philadelphia.

Lazarsfeld, P.F. (1935) 'The art of asking why'. *National Marketing Review*, **1**, 26–38.

Mosteller, F., Hyman, H., McCarthy, P.J., Marks, E.S. and Truman, D.B. (1949) *The Pre-election Polls of 1948: Report to the Committee on Analysis of Pre-election Polls and Forecasts*. US Social Science Research Council (Bulletin, no. 60), New York.

Oppenheim, A.N. (1966) *Questionnaire Design and Attitude Measurement*. Heinemann, London; Basic Books, New York.

Payne, S.L.B. (1951) *The Art of Asking Questions*. (Studies in Public Opinion, no. 3.) Princeton University Press, Princeton.

Smith, G.H. (1954) *Motivation Research in Advertising and Marketing*. McGraw-Hill, New York.

Stuart, A. (1952) 'Reading habits in three London boroughs'. *Journal of Documentation*, **8**, 33–49.

Whitfield, J.W. (1950) 'The imaginary questionnaire'. *Quarterly Journal of Experimental Psychology*, **2**, 76–87.