7

Constructing questionnaires

Questionnaires are the most common method of collecting survey data. This chapter outlines how to:

- construct individual survey questions that work;
- get the wording of questions right;
- use a variety of question formats;
- test whether your questions satisfy the principles of good question design;
- structure and design questionnaires from individual questions;
- modify questionnaires depending on the way in which they are to be administered;
- pilot test a questionnaire to see if it works.

SELECTING AREAS

Typically, when using questionnaires, it is difficult to go back to people to collect additional information. Therefore it is crucial to think ahead and anticipate what information will be needed to ensure that the relevant questions are asked.

There are a number of ways of working out which questions to ask. First, the research problem will affect which concepts need to be measured (see Chapter 3). Second, the indicators we devise for these concepts are crucial in determining which questions to ask (see Chapter 4). Third, our hunches about the mechanisms by which variables are linked or about factors which might explain relationships will require that certain questions be included (see Chapter 16). Fourth, the way data are to be analysed affects what information is needed: it is pointless collecting information which cannot be analysed and frustrating to discover that you do not have the necessary data for certain analysis. It is not possible to develop a questionnaire that can be analysed properly unless you first understand methods of analysis (see Chapters 12 to 17). Finally, the method by which the questionnaire is to be administered affects the type of questions that can be asked. If the questionnaire is administered by a trained interviewer more complex questions can be used since there is opportunity for clarification. In addition, follow-up questions which draw on answers to earlier questions can be used. With self-administered questionnaires such as those sent out by post you need to concentrate on clarity and simplicity.

In summary, the art of questionnaire design involves thinking ahead about the research problem, what the concepts mean and how we will analyse the data. The questionnaire should reflect both theoretical thinking and an understanding of data analysis.

BOX 7.1
Checklist to guide questionnaire content in explanatory research

Do you have questions that provide:

1 measures of the dependent variable(s);
2 measures of the independent variable(s);
3 measures of test variable;
4 background measures.

QUESTION CONTENT

It is helpful to distinguish between five distinct types of question content: behaviour, beliefs, knowledge, attitudes and attributes (Dillman, 1978: 80). Imagine that you are conducting a study on the topic of workforce participation of mothers of preschool age children and you had a sample of mothers—some with young children, others with older children. Before you could formulate any questionnaire items you would need to be very clear about the issues in which you are interested.

If you were interested in *behaviour* you would formulate questions to establish what people *do*. For example, you would ask whether the respondent is working or did work with a preschool age child. Depending on the precise research question this can provide useful information. It can provide a map of which types of mothers work and which types do not and may help locate factors which facilitate or hinder workforce participation. But too often researchers try to use behavioural measures to extrapolate to beliefs and attitudes. This is open to real dangers of misinterpretation. Since people are neither very consistent nor rational and may not have the luxury of behaving as they might like, any conclusions we can draw about beliefs or attitudes from behaviour are very limited.

If you are interested in *beliefs*—in what people believe is *true or false*—you need to ask quite different types of question. For example, you might ask people for their estimate of the percentage of mothers with preschool-age children who are in the paid labour force or ask about what they believe to be the effects of day care centres on the emotional development of preschool-age children. The focus of belief questions is on establishing what people think is true rather than on the accuracy of their beliefs.

Knowledge questions have some similarities to belief questions. They seek to discover respondents' knowledge of particular facts, such as the percentage of children in child-care, the government programs designed to assist parents with pre-schoolers to work part-time, and the tax implications of working part-time. However, while belief questions are designed to discover *what* people believe, knowledge questions are formulated to establish the *accuracy* of their beliefs. The difference between knowledge and belief questions lies less in the construction of the question than in the way the answers are interpreted and used.

Belief questions can be distinguished from those that aim to establish the respondent's *attitudes*. Whereas belief questions ascertain what the respondent thinks is true, attitude questions try to establish what they think is *desirable*. An attitudinal focus might ask about attitudes regarding whether or not mothers with preschool-age children ought to participate in the workforce.

Finally, *attribute* questions are designed to obtain information about the respondents' characteristics. Such questions would normally include information about their age, education, occupation, gender, ethnicity, marital status and so forth. For the study of workforce participation of mothers with preschool-age children you might ask about attributes such as the number of children, the age of the child, income, type of job, whether the job was full-time or part-time and other related information.

It is important to be clear about the precise type of information required for a number of reasons. First, the failure to adequately distinguish between these five types of information arises from a lack of clarity about the research question and inadequate conceptualisation. This can lead to the collection of quite the wrong type of information. If we are interested in exploring people's actual behaviour, a set of questions that in fact taps beliefs or attitudes will be of little use. Second, we might be interested in all five types of information. An awareness of the five types of information that can be collected should lead to the systematic development of questions for each type rather than a haphazard set of questions on the broad topic which may or may not tap all types of topic. Third, when it comes to analysis and the development of scales (see Chapter 11), it is important to develop

composite measures; however, these normally need to be composite measures of the same type of information. Attitude questions can be combined with other attitude questions to form an index of some sort or another but it would normally be quite inappropriate to combine the four types of information into a single measure: they tap quite different things.

Direction, extremity and intensity of attitudes

It is important to distinguish the *direction* of a person's attitude from both the *extremity* of their position and the *intensity* with which they hold that position. Each of these three aspects of an attitude requires a different sort of question.

We may wish to know people's views about whether government economic policy ought to be directed more at reducing inflation and the government deficit or at reducing unemployment. We could discover the *direction* of a person's attitude by simply asking them which of the two policy directions they think is the more desirable. But we could learn more by asking how *extreme* their view is. There are two ways in which this is commonly done. The first is to provide a statement that expresses a *position* (e.g. the government's first priority ought to be to reduce unemployment even if this leads to increased inflation and problems with the deficit) and asking them to say how strongly they agree or disagree with it. Alternatively, a seven-point scale might be used in which 'reduce inflation' is placed at one end and 'reduce unemployment' at the other. Respondents can be asked to indicate where they would place themselves between these two positions.

This approach does not necessarily detect the *intensity* with which a position is held. Although extremity and intensity of an attitude may often go together, they are not the same thing. A person can hold extreme positions but do so with little passion. People may vote for extreme left- or right-wing political parties without having a fervent commitment to that party. Questions that measure a person's attitude *position* can usefully be followed up with questions to detect the attitude *intensity*.

PRINCIPLES OF QUESTION DESIGN

Before dealing with the specifics of question wording and answer formats it is important to highlight six broad principles that must be built into question design.

Reliability

The question should be answered in the same way on different occasions if given to the same person (assuming that the person has not changed in the meantime). A question that fails to achieve consistent responses is unreliable. Ambiguous or vague question wording may produce unreliable responses as respondents 'read' the question differently on different occasions.

Validity

A valid question is one that measures what we think it does. Thus if we use self-rated health (i.e. how healthy are you?) as a measure of health we need to be confident that it measures health rather than something else such as optimism and pessimism. If we use an IQ test to measure intelligence we want to be sure that it does in fact measure intelligence, rather than social class background.

Discrimination

Explanatory survey analysis requires that there is variation in the sample on the key variables. If we wanted to see if there was a link between gender and income we would have to have a sample in which there were men and women and in which there was a good variety of income levels. Low variance may reflect real homogeneity (or sameness) within the sample (e.g. gender and income level). But it can also stem from poor question design—a limited range of response alternatives can produce low variance. If our questions do not pick up actual variation in the sample then the information obtained by the question will be of no use at the data analysis stage. For example, if we asked about income and simply had two response categories of 'Less than $100 000 a year' and 'Over $100 000 a year' we would not identify much variation in the sample. Most people would be in the 'Less than $100 000 a year' category. For analytic purposes all these people would be treated as though they were the same despite what might be considerable income differences. A question with finer grained response categories would identify greater variation across the sample. Good questions

will be sensitive to measuring real and meaningful differences in a sample.

When measuring attitudes low variance can also stem from using extreme attitude statements. For example, an extreme statement such as 'In a country such as this, assassination of political leaders is acceptable in order to bring about political change' will probably yield low variance. The solution is to provide sufficient response alternatives to detect meaningful variation and to avoid using extreme or absolute statements.

Response rate

Ideally all respondents will answer every question that applies to them but experience tells us that some questions can elicit relatively high non-response. Non-response needs to be minimised both because of the loss of information and the data analysis difficulties it introduces. Non-response is affected by question content, question construction, method of administration and questionnaire length. Intrusive, sensitive, irrelevant or repetitive questions as well as those that are poorly worded, difficult to understand, difficult to answer or have insufficient response categories can frustrate respondents and produce non-response.

Same meaning for all respondents

When analysing questionnaires we assume that all respondents have answered the same questions. However if respondents interpret the questions in different ways they are effectively answering different questions. If I use the term 'old people' or 'elderly people' in a question respondents will use different definitions for the terms 'old' or 'elderly' and in effect be answering different questions. Careful question design is needed to minimise this problem.

Relevance

Each question must earn its place in your survey. For each question ask yourself whether it really is necessary.

WORDING QUESTIONS

Considerable attention must be given to developing clear, unambiguous and useful questions. To do this the wording of the questions is fundamental. The checklist of seventeen questions in Box 7.2 will help you to avoid the most obvious problems with question wording.

1 Is the language simple?
Avoid jargon and technical terms. Look for simple words without making questions sound condescending. Use simple writing guides or a thesaurus to help (see Gowers, 1962; Strunk and White, 1972). A question such as 'Is your household run on matriarchal or patriarchal lines?' will not do!

2 Can the question be shortened?
The shorter the question the less confusing and ambiguous it will be. Avoid questions such as: 'Has it

BOX 7.2 ***Question wording checklist***

1 Is the language simple?
2 Can the question be shortened?
3 Is the question double-barrelled?
4 Is the question leading?
5 Is the question negative?
6 Is the respondent likely to have the necessary knowledge?
7 Will the words have the same meaning for everyone?
8 Is there a prestige bias?
9 Is the question ambiguous?
10 Is the question too precise?
11 Is the frame of reference for the question sufficiently clear?
12 Does the question artificially create opinions?
13 Is personal or impersonal wording preferable?
14 Is the question wording unnecessarily detailed or objectionable?
15 Does the question have dangling alternatives?
16 Does the question contain gratuitous qualifiers?
17 Is the question a 'dead giveaway'?

happened to you that over a long period of time, when you neither practised abstinence nor used birth control, you did not conceive?'.

3 Is the question double-barrelled?

Double-barrelled questions are those which ask more than one question. The question 'how often do you visit your parents?' is double-barrelled. Separate questions about a person's mother and father should be asked.

4 Is the question leading?

A leading question is one where either the question structure or wording pushes people to provide a response that they would not have given had the question been asked in a more neutral way (Payne, 1951; Bradburn and Sudman, 1979; Belson, 1981). Questions such as 'Do you oppose or favour cutting defence spending even if cuts turn the country over to communists?' are obviously leading. Leading questions give respondents the impression that there is a 'correct' response. Avoid linking an attitude position, policy or whatever with a prestigious person. Avoid phrases such as 'Do you agree that . . .?' or 'Does this seem like a good idea to you?'. The particular terminology you use can be leading. Think of the different impact of the choice of words 'abortion', 'killing unborn babies' or 'ending a pregnancy'.

5 Is the question negative?

Questions which use 'not' can be difficult to understand—especially when asking someone to indicate whether they agree or disagree. The following question could be confusing:

> Marijuana should not be decriminalised
> –Agree
> –Disagree

Rewording the question to 'Marijuana use should remain illegal' avoids the confusion caused by using 'not'.

6 Is the respondent likely to have the necessary knowledge?

When asking about certain issues it is important that respondents are likely to have knowledge about the issue. A question which asks 'Do you agree or disagree with the government's policy on legalising drug injecting rooms?' would be unsatisfactory. For issues where there is doubt, we might first ask a filter question to see if people are aware of the government's policy on drug injecting rooms and then ask the substantive question only if people answered 'yes' to the filter question. Alternatively, we should offer the respondent the opportunity to say that they are not sure what the government's policy is.

7 Will the words have the same meaning for everyone?

Depending on factors such as age group, subcultural group and region, the meaning of some words will vary, so care must be taken either to avoid such words or to make your meaning clear. People also vary in how they define certain terms. For example, the answers people give to a question that asks them if they have been a victim of a crime in the last five years will depend on what they include in their definition of crime. For example, despite its illegality, some people may exclude domestic violence from their definitions of crime, thus leading to its underreporting.

8 Is there a prestige bias?

When an opinion is attached to the name of a prestigious person and the respondent is then asked to express their own view on the same matter, the question can suffer from prestige bias. That is, the prestige of the person who holds the view may influence the way respondents answer the question. For example, 'What is your view about the Pope's policy on birth control?' could suffer from prestige bias. Effectively the question is double-barrelled: the answer may reflect an attitude about the Pope or about birth control—we cannot be sure which.

9 Is the question ambiguous?

Ambiguity can arise from poor sentence structure, using words with several different meanings, use of negatives and double negatives, and using double-barrelled questions. The best way to avoid ambiguity is to use short, crisp, simple questions.

10 Is the question too precise?

While we need to avoid questions which invite vague and highly imprecise responses we also need to avoid requiring answers that need more precision than people are likely to be able to provide reliably. Precise answers are not necessarily accurate answers. Asking for too precise an answer can produce

unreliable responses and add nothing useful to the study (Foddy, 1993). For example, asking people 'How many times in the last year did any member of your household visit a doctor?' may yield precise figures but they are likely to be both inaccurate and unreliable.

11 Is the frame of reference for the question sufficiently clear?

If you ask 'How often do you see your mother?', establish within what time frame—within the last year? the last month? If you mean the frequency within the last year, ask 'Within the last year how often would you have seen your mother on average?' and then provide alternatives such as 'daily' through to 'never' to help further specify the meaning of the question.

12 Does the question artificially create opinions?

On certain issues people will have no opinion. You should therefore offer people the option of responding 'don't know', or 'no opinion'. This can lead to some people giving these responses to most questions which can create its own problems, but not including these alternatives will produce highly unreliable, and therefore useless, responses (see p. 106 for further discussion).

13 Is personal or impersonal wording preferable?

Personal wording asks respondents to indicate how 'they' feel about something, whereas the impersonal approach asks respondents to indicate how 'people' feel about something. The approach you use depends on what you want to do with the answers. The impersonal approach does not provide a measure of someone's attitudes but rather the respondent's perception of other people's attitudes.

14 Is the question wording unnecessarily detailed or objectionable?

Questions about precise age or income can create problems. Since we normally do not need precise data on these issues we can diffuse this problem by asking people to put themselves in categories such as age or income groups.

15 Does the question have dangling alternatives?

A question such as 'Would you say that it is frequently, sometimes, rarely or never that . . .?' is an awkward construction. The alternative answers are provided before the respondent has any subject matter to anchor them to. The subject matter should come *before* alternative answers are listed.

16 Does the question contain gratuitous qualifiers?

The italicised qualifiers in the following examples would clearly affect the way people answer the question—they effectively present an argument for a particular response. 'Do you oppose or favour cutting defence expenditure *even if it endangers our national security*?' and 'Do you favour or oppose increasing the number of university places for students *even if it leads to a decline in standards*?'

17 Is the question a 'dead giveaway'?

Absolute, all-inclusive or exclusive words are normally best avoided. Examples of such 'dead giveaway' words (Payne, 1951) are: all, always, each, every, everybody, never, nobody, none, nothing. Since these words allow no exceptions few people will agree with the statement that includes them and this in turn will result in low variance and poor question discrimination.

SELECTING QUESTION TYPE

The other aspect of question construction is to decide on the response format. Should it be open or closed? If a closed format is used then a number of alternative types are available.

Open and closed formats

A closed or forced-choice question is one in which a number of alternative answers are provided from which respondents are to select one or more of the answers. An open-ended question is one for which respondents formulate their own answers.

There is disagreement about which style is preferable. A major problem of forced-choice questions is that on some issues they can create false opinions either by giving an insufficient range of alternatives from which to choose or by prompting people with 'acceptable' answers. Further, the forced-choice approach is not very good at taking into account people's qualifiers to the answers they select.

There are, however, a number of advantages to *well-developed* forced-choice questions. Where the

questionnaire is long or people's motivation to answer is not high, forced-choice questions are useful since they are quick to answer. This is particularly so if the questionnaire is self-administered rather than administered by a skilled interviewer who can establish rapport and increase motivation.

From a researcher's point of view, forced-choice questions are easier to code (see Chapter 9). Answers to both closed and open questions need to be grouped into categories at some stage. The difficulties of doing this with open-ended questions often mean that they never get used. Even if they are grouped, researchers normally interpret answers and put them in categories. Researchers can misinterpret the answers and thus misclassify responses. Forced-choice questions allow respondents to classify themselves, thus avoiding coders misclassifying what people meant.

A further advantage of closed questions is that they do not discriminate against the less talkative and less articulate respondents. Asking people to formulate their own responses is fine for those who can do it but the danger is that researchers will be overly influenced by these responses and ignore the opinions of the less articulate and less fluent.

A set of alternative responses can serve as useful prompts for respondents. For example, a question asking about the newspapers and magazines a person has read in the last week will detect a higher readership level if the names of newspapers and magazines are listed in a checklist than if the open-ended question is simply asked without the list of responses.

If forced-choice questions are used, it is necessary to put a lot of thought into developing alternative responses. The range must be exhaustive: a thorough range of responses must be listed to avoid biasing responses. This can be done by careful pilot testing using less structured approaches to locate the range of likely responses and by using the category called 'other (please specify)' to allow for unanticipated responses.

The choice of open or closed questions depends on many factors such as the question content, respondent motivation, method of administration, type of respondents, access to skilled coders to code open-ended responses and the amount of time available to develop a good set of unbiased responses. There is no right or wrong approach.

RESPONSE FORMATS AND LEVEL OF MEASUREMENT

In Chapter 12 the concept of level of measurement will be discussed. Go to pp. 203–6 and read the section on levels of measurement. In that section you will see that the level of measurement of a variable is based on the response categories of a variable and the relation of one response category to another.

The level of measurement of a variable is fundamental in the choice of statistical methods when we come to analyse the data. The way we frame many questions will influence the level of measurement of a variable and thus the way we analyse data later on.

From a data analysis perspective it is generally best to have data that are measured at the interval level. This enables us to use a wider range of statistical methods and allows us to use the more powerful techniques should we need them. From this perspective it is desirable to design questions so that they result in interval-level variables. However, the principles of good questionnaire design are not always consistent with this and we may have to settle for less precise question formats that collect ordinal-level data. We will then be restricted to methods of analysis that only require this sort of data.

We have some control over the level of measurement of a variable. The way we ask the question, or more precisely, the sort of response alternatives provided will yield data at a nominal, ordinal or interval level (see Figure 7.1).

DEVELOPING QUESTION RESPONSES

There are three guiding principles when developing question responses.

Exhaustiveness (or inclusiveness)

Ensure that the response alternatives provide a sufficient range of responses to cover all respondents. A question that asks about marital status and includes only 'married' and 'divorced' as alternatives is not providing an exhaustive set of options.

Numeric rating scales (see below) are a good way of providing an exhaustive set of responses for many questions. Attitude questions should generally include a 'don't know' or 'no opinion' option (see p. 106) so that those with no opinion are provided for. For some

Interval level

How many years of formal education have you completed since you left secondary school? (circle the number that applies to you)

0 1 2 3 4 5 6 7 8 9 9+

Ordinal

What is the highest level of qualification you have completed since leaving secondary school?

☐ Certificate
☐ Diploma
☐ Bachelor's degree
☐ Graduate diploma
☐ Masters degree
☐ PhD

Nominal

Since leaving secondary school which of the following best describes what you have been doing?

☐ Further study
☐ Working full-time (for an income)
☐ Working and studying
☐ Home duties
☐ Travelling
☐ None of the above

Figure 7.1 Question format and level of measurement

questions it is appropriate to add an open category where respondents can create their own answer if the set provided has not been exhaustive. For example, a question asking about the respondents' country of birth might provide a list of the most common countries but add a final, catch-all category of 'Other (please specify)' to cover those respondents not covered by the preset responses.

Exclusiveness

This principle means that for each 'question' a person can provide one and *only one* answer to the 'question'. That is, the alternate responses should be mutually exclusive. This is not a problem with rating scales and questions where the response alternatives are graded along a single continuum. Respondents may have difficulty identifying where on the continuum they lie but in principle they lie at a particular point.

Exclusiveness is a problem where a person might quite legitimately select more than one of the alternative responses, as illustrated in Figure 7.2.

There are several ways of dealing with this problem. The first would be to add additional categories such as 'Both' or 'Neither'. Another solution is to reduce the choices to a single choice by asking people to nominate which of the alternatives is *most* important. In this case the question could be rephrased to read 'In your workplace which would you say is the more important for getting ahead: merit or gender?'

In your workplace would you say that getting ahead is based on merit or gender?

☐ Merit
☐ Gender

Figure 7. 2 Non-exclusive responses

Another solution is to treat each response as a separate question or variable. For example, instead of asking whether gender or merit is the more important we could ask respondents to show how important each factor is (e.g. 'Important', 'Not important', and 'No opinion'). Similarly, in checklist questions each item can be thought of as a separate variable for data analysis with the categories of 'selected' and 'not selected' (see the section on multiple response coding in Chapter 9). In this way the principle of exclusiveness is preserved.

Balancing categories

An imbalance of responses in one direction can produce bias (Payne, 1951). Where response categories can be ordered from high to low there should be the same number of response alternatives either side of what might be considered the neutral position. For example, the alternatives in Figure 7.3 will introduce bias and underestimate the level of disapproval.

DEVELOPING RESPONSE ALTERNATIVES FOR CLOSED-CHOICE QUESTIONS

A range of ways of responding to closed-choice questions are available. The type of response

☐ Completely approve
☐ Strongly approve
☐ Approve
☐ Neither approve nor disapprove
☐ Disapprove

Figure 7.3 Unbalanced response alternatives

alternative has major implications for response rates to questions, coding and the way data are analysed. The particular formatting of these response alternatives will differ depending on the mode of questionnaire administration (see Chapter 8).

Numerical rating scales

Rating scales involve a set of responses where the alternative answers are ordered from low to high. Respondents need to indicate where between the low and high extremes lies their attitude. There are a variety of ways in which rating scales are structured.

Likert scales

This approach to measuring attitudes involves providing a statement that reflects a particular attitude or opinion. Respondents indicate their level of agreement or disagreement with the statement. Usually respondents are given the alternatives of strongly agree, agree, neither agree nor disagree, disagree and strongly disagree. This style of question may be presented as a single item question or as a set of questions arranged in a grid format. The grid format, apart from saving space, is easy to answer and is used for sets of items that form scales (see Chapter 11). For the purpose of data analysis each statement to which an answer is sought is a separate variable (see Figure 7.4).

Horizontal rating scales

These scales provide respondents with opposite attitude positions and asks them to indicate with a number where, between the positions, their own view falls (see Figure 7.5).

Semantic differential

This method provides *opposite adjectives*, rather than attitude positions, to describe someone or something. These adjectives are placed at the opposite ends of the numeric scale. Each pair of adjectives provides a separate variable for data analysis (see Figure 7.6).

Vertical rating ladder

You might want to ask people to rate the status of particular universities such as Harvard, Oxford, Princeton, Cambridge and Melbourne. You might provide respondents with a rating ladder like that below and ask them to indicate where on this ladder they would place the status of each university. Different universities could share the same rating and each university would be treated as a separate variable for analysis purposes (see Figure 7.7).

Scores

Rather than using graphical rating scales we might simply ask the respondent to write in a score to

To what extent do you agree or disagree with each of the statements as far as your immediate supervisor is concerned?	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
1 Criticises people in a manner which builds their motivation	☐	☐	☐	☐	☐
2 Admits to their mistakes	☐	☐	☐	☐	☐
3 Takes action without waiting to be asked to	☐	☐	☐	☐	☐
4 Praises others' ideas and contributions	☐	☐	☐	☐	☐
5 Takes personal responsibility when things go wrong	☐	☐	☐	☐	☐

Figure 7.4 Likert-style questions in a grid format

Some people think that the government should provide for proper care of the elderly while others think that it is the responsibility of families.

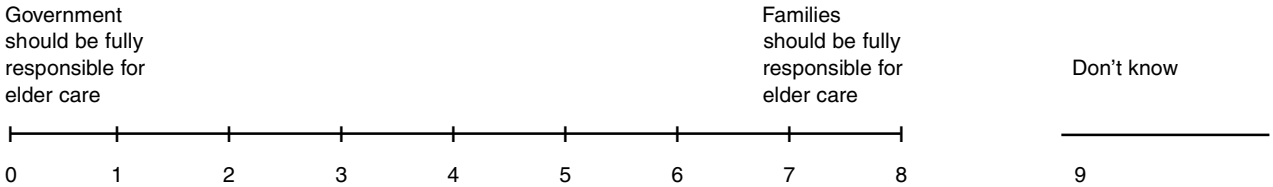


Figure 7.5 Horizontal rating scale

I would like you to describe your workplace using the following set of descriptions. For each description circle the number below the line to indicate where your workplace falls.

Figure 7.6 Semantic differential rating scales

	Rank	University
High	10	
	9	Oxford; Cambridge
	8	Harvard
	7	
	6	Princeton
	5	
	4	
	3	Melbourne
	2	
Low	1	

Figure 7.7 Vertical rating ladder

indicate their rating. There are a number of formats for doing this.

Out of 10

Instead of providing the diagram you could simply ask respondents to indicate their rating of something by giving it a score out of 10. For example, the question in Figure 7.8 might be used to study job satisfaction.

Feeling thermometer

With the feeling thermometer (Figure 7.9) each group becomes a variable during the analysis.

The list below describes various features of jobs. For each feature think of your current job and say how satisfied you are with that aspect of your job by giving a score out of 10. If you are completely satisfied you should give this feature a score of 10. If you are completely dissatisfied give it a score of 0. If you are neither satisfied nor dissatisfied give it a score of 5. You can give any score between 0 and 10.
How would you rate your current job in relation to:

☐ ☐ the feeling of accomplishment it gives you
☐ ☐ the level of enjoyment and interest it gives
☐ ☐ the security and predictability it gives you
☐ ☐ the people you get to meet at work
☐ ☐ the amount of money you earn

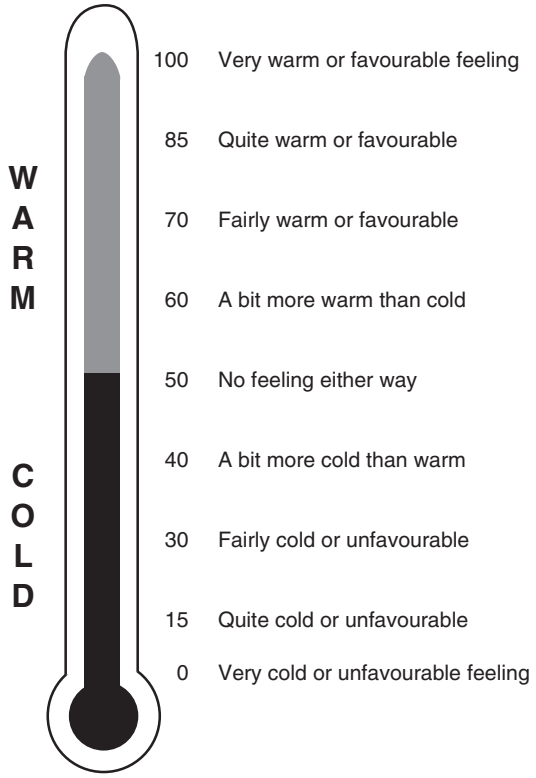
Figure 7.8 Score out of 10 rating scale

Despite the differences between these various rating scales they have important characteristics in common.

- They all require that respondents give one and only one response to each item.
- They all produce variables where the responses can be ordered from high to low.
- The way in which each item is answered is not constrained by the way in which other items in a set are answered. This is distinctly different from questions that require the ranking of responses.

Ranking

The ranking format requires respondents to rate the importance or strength of agreement *relative to the way other items in the set have been rated*. This format



We'd like to get your feelings about some jobs. We'd like you to rate each job with what we call a feeling thermometer. Ratings of between 50° and 100° mean that you feel the job is very desirable. Ratings between 0° and 50° mean that you don't care too much for that job. Place your rating in the boxes next to the job.

☐☐☐ Computer programmer
☐☐☐ Medical specialist
☐☐☐ Management consultant
☐☐☐ Lawyer
☐☐☐ Accountant
☐☐☐ Scientist
☐☐☐ Academic
☐☐☐ Engineer

Figure 7.9 Feeling thermometer

provides answers that indicate the *relative* rather than the *absolute* importance of items. Figure 7.10 provides an example.

Depending on how many items are listed we might just ask that the top two and bottom two items be ranked. Where there are relatively few items (e.g. four or five) we might ask that they all be ranked. With this format there are two ways of creating variables for analysis (see Chapter 14).

Listed below is a set of issues that can influence the way in which people decide to vote in general elections. Please rank each of these issues to indicate how important they are to *you* when you decide to vote. Place 1 in the box next to the most important issue, 2 next to the second most important issue and so on. Do not place the same number in more than one box.

☐ Policies to reduce unemployment
☐ Improving the environment
☐ Spending more money on education
☐ Getting tough on crime
☐ Reducing taxation
☐ Improving social welfare support
☐ Improving health services
☐ Reducing immigration

Figure 7.10 Ranking response format

Checklists

This format involves listing a set of items and asking that the respondent simply select those that apply (see Figure 7.11).

Binary choice formats

These formats require the respondent to choose between one of two fixed alternatives.

Dichotomous questions

These questions ask the respondent to select between one of two alternatives. Checklists are effectively binary choice questions (select or do not select). Other examples are illustrated in Figure 7.12.

Paired comparisons

Another form of the binary choice is a set of paired comparisons where the respondent is given a set of

Listed below are some adjectives, some of which are 'favourable', some of which are 'unfavourable', some of which are neither. Please tick the boxes beside the characteristics that best describe you as a person. Most people choose three or four, but you may choose more or fewer if you want.

☐ Ambitious
☐ Athletic
☐ Cautious
☐ Good looking
☐ Moody
☐ Happy
☐ Obliging
☐ Highly strung
☐ Poised

Figure 7.11 Checklist response format

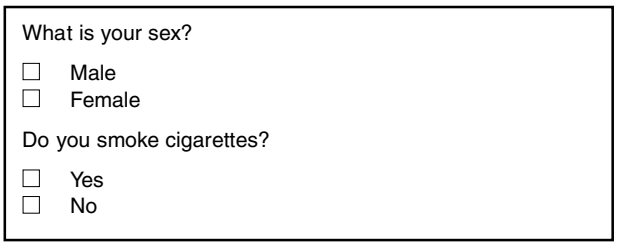


Figure 7.12 Dichotomous response format

overlapping pairs of items and asked to select one response from each pair (see Figure 7.13).

Multiple choice formats

This format requires respondents to choose just *one* response from a list of three or more alternatives. Many of the numeric rating scales are actually a form of multiple choice format. Additional types of multiple choice formats are described below.

Choice between multiple nominal categories
Respondents are asked to select one alternative from a list of responses. These responses have no set order and cannot be ranked in any sense from high to low. An example is a question on marital status (see Figure 7.14).

Choice between ordinal categories
Other multiple choice questions will have a set of responses that should be ordered from low to high (see Figure 7.15).

Choice between ordered attitude statements
While the Likert format asks respondents for the extent of agreement with a particular attitude statement this format provides alternative attitude positions and asks which is closest to the respondents' own view (see Figure 7.16).

Numerical answers
Some questions can be answered by a numerical response. This might be a precise numeric answer as required in example (a) in Figure 7.17 or it may group a set of numeric answers (e.g. income in dollars) into broader numeric bands as illustrated in example (b) in Figure 7.17.

NON-COMMITTAL RESPONSES

No opinion and don't know responses

Generally you should allow for a 'don't know' or 'no opinion' response. There are many issues to which people will have given no thought and hold no

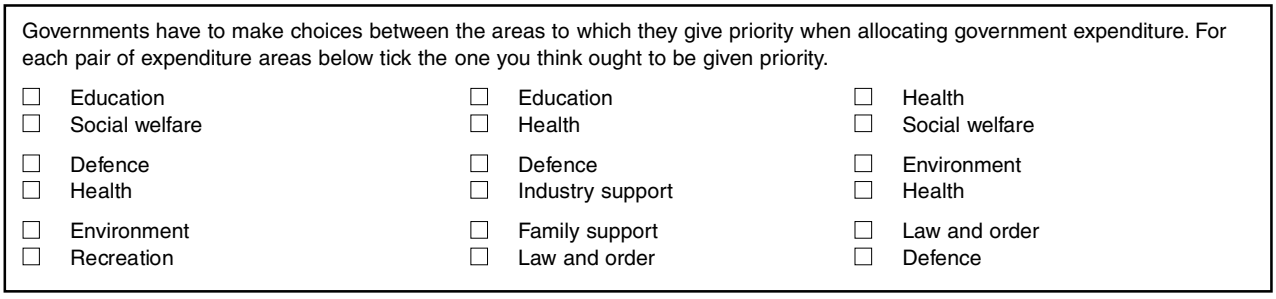


Figure 7.13 Paired comparison response format

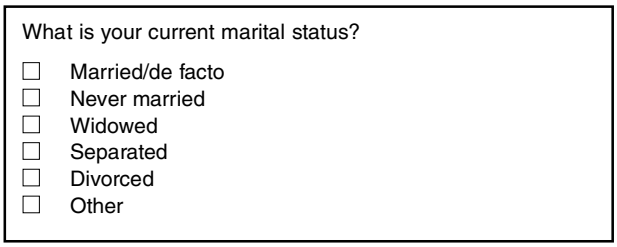


Figure 7.14 Multiple nominal responses

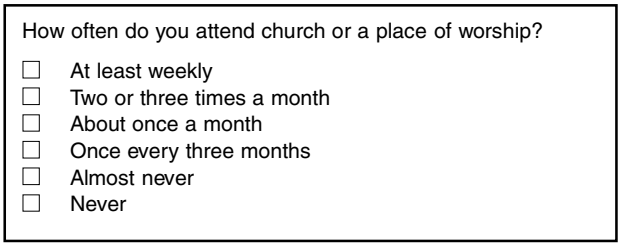


Figure 7.15 Multiple ordinal responses

Which of the following statements comes closest to your belief about God (tick one box only):

☐ I know God really exists and have no doubts about it.
☐ While I have doubts, I feel that I do believe in God.
☐ I find myself believing in God some of the time but not at others.
☐ I don't believe in a personal God, but I do believe in a higher power of some kind.
☐ I don't know whether there is a God and I don't believe there is a way to find out.
☐ I don't believe in God.

Figure 7.16 Multiple attitude statements

a How many people under the age of eighteen live in your household on a regular basis?
☐ ☐

b Questions with numerical answers can require precise numeric answers or may place numeric answers into groups. Instead of asking for precise income we might ask respondents to nominate to which income band they belong. For example:

What was your own income from your salary or wage, before tax, in the year 2001?

☐ None
☐ $1–$4999
☐ $5000–$9999
☐ $10 000–$14 999
☐ $15 000–$19 999
☐ $20 000–$29 999
☐ $30 000–$39 999
☐ $40 000–$49 999
☐ $50 000–$64 999
☐ $65 000–$79 999
☐ $80 000–$99 999
☐ $100 000 or more

Figure 7.17 Numeric response format

opinion. To force them to express an opinion where they really do not have one is to create false and unreliable answers (Converse and Presser, 1986).

The danger with using 'don't know' and 'no opinion' alternatives is that some respondents select them out of *laziness*. We can discourage respondents routinely selecting these responses by making them less conspicuous in the questionnaire (Schuman and Presser, 1981; Converse and Presser, 1986). In written self-administered questionnaires there is little alternative but to include 'don't know' or 'no opinion' along with other alternatives. However with telephone and face-to-face interviews we can rely on respondent-initiated non-response rather than offering it as an up-front alternative. That is, we will accept 'don't know' without *offering* it as an alternative (see Figure 7.18).

Inclusion of the middle alternative

Another sort of non-committal answer is the 'sitting on the fence' answer. That is, where there is an ordered set of responses (e.g. for and against something) should we provide a middle alternative (neither for nor against)? There is some disagreement about what to do in this situation. Some people argue that the middle alternative should *not* be included since omitting it forces people to indicate the direction of their opinion and stops people sitting on the fence. On the other hand, including the middle position avoids artificially creating a directional opinion (see Converse and Presser, 1986; Sudman and Bradburn, 1982; Presser and Schuman, 1980).

One way of proceeding is not to *offer* the middle alternative when the questionnaire is administered by an interviewer but to record a middle position response if it is *volunteered*. In self-administered questionnaires it is desirable to offer the middle position to avoid forcing people to indicate a level of commitment that they do not have.

NUMBER OF RESPONSE CATEGORIES

Experts do not agree about the number of response categories that should be provided (Schwarz et al., 1985).

Dichotomies: One approach is to ask the respondent to select between one of two alternatives (Payne, 1951). For example, we might ask customers if they are satisfied or dissatisfied with the level of service they received. The problem with dichotomous

Do you agree or disagree that all people ought to have free access to government funded health care?

☐ Agree
☐ Disagree
☐ Don't know *(instruction to interviewer—do not offer this response; code only if respondent will not offer an agree or disagree response)*

Figure 7.18 Respondent-initiated 'don't know' response

responses is that they often provide a poor response distribution because people's real position lies somewhere between the two extremes (e.g. somewhat satisfied).

Five point scales: The Likert format (see Figure 7.4) provides five response alternatives which gives more flexibility. It provides a measure of intensity, extremity and direction. If needed you can later collapse five categories down to two or three.

Longer scales: The use of longer scales can have some advantages in that they allow for greater discrimination. For example, a ten-point scale allows for the detection of finer differences between people than would be possible with a five-point scale (see Alwin, 1997). This can be useful for attitudes where people tend to only use the top end of the scale—a common problem with questions that measure satisfaction. For example, questions that ask about life satisfaction typically indicate high levels of satisfaction. This partly reflects the social desirability of certain responses but a ten-point scale allows people to give a 'satisfied' response while still indicating some qualification to this.

However too many fine distinctions can be baffling and there is the danger that fine distinctions confuse greater precision with greater accuracy. In deciding on the number of response categories it is helpful if you have some sense of the spread of the variable when fewer categories are used. The main justification for using a larger number of response categories is that fewer categories are insensitive to real differences. The other consideration is sample size and the way data will be analysed. There is probably little point in using nine response categories when in the final analysis the categories will be collapsed down to three for analysis purposes.

RESPONSE SETS

Some respondents are liable to provide a certain type of answer regardless of the content of the question. There is the problem of *acquiescence*—the tendency to *agree* with a statement regardless of its content—and the problem of *social desirability*—the tendency to provide the respectable rather than the true response.

Social desirability

Many people answer survey questions so that they look good in their own eyes and in the eyes of interviewers. Consequently socially 'desirable' behaviours (e.g. amount of exercise) are over-reported while socially 'undesirable' behaviours and attitudes (e.g. alcohol consumption, sexist and racist attitudes) are under-reported (Bradburn and Sudman, 1978; Sudman and Bradburn, 1982; Foddy, 1993).

The techniques listed in Box 7.3 can help reduce social desirability as a factor in question responses.

Acquiescent response sets

Acquiescence (Foddy, 1993) is greatest among respondents with low education, in face-to-face interviews, where general rather than specific questions are used and where respondents have not really formed an opinion. It is more likely in attitude questions that use the agree–disagree format than in rating scales or the selection between different attitude statements.

Where acquiescence is likely to be a problem adopt a response format that minimises the problem and make sparing use of the agree–disagree format. Where the agree–disagree format is needed ensure

WEB POINTER 7.1 ***Sets of response alternatives***

The sites below provide a very useful set of response wordings with a varying number of response categories for questions where the response categories can be ranked from high to low in some respect.

Sample Sets of Response Alternatives	www.au.af.mil/au/hq/selc/smpl-h.htm
The intensity of words	www.au.af.mil/au/hq/selc/smpl-g.htm

Visit www.social-research.org to use these links and to check for updates and additions.

BOX 7.3
Reducing social desirability response sets

There are four main techniques of asking questions to reduce social desirability problems:

1 everybody does it (question a);
2 use an authority (question b);
3 build in an excuse (question c);
4 ask a less specific question (question d).

a Even the calmest of parents get angry at their children some of the time. Did your children do anything in the last seven days to make you feel angry?

b Many doctors now think that drinking wine reduces heart attacks and aids digestion. Have you drunk any wine in the last seven days?

c We know that people are often very busy and can find it difficult to find time to engage in regular exercise. How often have you engaged in exercise designed to improve your fitness in the last seven days?

d Have you ever, *even once*, hit your partner in anger?

that the 'pro' attitude statements are matched by a similar number of 'anti' statements (e.g. a set of attitude questions designed to measure conservatism should aim for a similar number of conservative and non-conservative statements). In this way if people are inclined to agree with any statement the acquiescence effect should cancel out.

QUESTIONNAIRE LAYOUT

There are six areas to which attention needs to be given when combining questions into a questionnaire.

Answering procedures

With open-ended questions ensure that you leave sufficient space for answers to avoid having to cram in responses. But do not leave so much space as to discourage completing the questionnaire because the task appears too daunting.

With closed questions respondents can be asked to either tick appropriate boxes or brackets or circle a number next to responses (see Figure 7.19).

When using any of these procedures, the area for answering can be on the left or right of the response but make sure you justify your typing on the answer side as below.

1 [ ] Agree		Agree [ ] 1
2 [ ] Disagree	OR	Disagree [ ] 2
3 [ ] Can't decide		Can't decide [ ] 3

Electronic means of administering questionnaires (see Chapter 8) have enabled an additional range of ways of respondents answering questions (see Web Pointer 7.3).

Contingency questions

Since you do not want respondents to waste time reading questions which are not relevant to them we

WEB POINTER 7.2 ***Three web-based questionnaire construction guides***

Good general article on designing questionnaires and questions.	www.css.edu/users/dswenson/web/question.htm
Brief guide to questionnaire construction.	www.ericae.net/ft/tamu/vpiques3.htm
Very basic guide to questionnaire construction.	www.webcom.com/ygourven/quest12.html

Visit www.social-research.org to use these links and to check for updates and additions.

Square brackets, parentheses or boxes (tick the box)

1 [ ] Agree		1 ( ) Agree		1 ☐ Agree	
2 [ ] Disagree	OR	2 ( ) Disagree	OR	2 ☐ Disagree	
3 [ ] Can't decide		3 ( ) Can't decide		3 ☐ Can't decide	

Precoding (circle the number)
1 Agree
2 Disagree
3 Can't decide

Figure 7.19 Different answering formats for closed-choice questions

WEB POINTER 7.3 ***Computer-based answering formats***

Computer-based surveys have introduced new ways of formatting questions that can be illustrated online. Go to the following internet address and select the 'respondent's view' for each of the question types provided:

www.surveysaid.com/marketing_masters/ssdocs/screens.htm

Visit www.social-research.org to use these links and to check for updates and additions.

can use filter or contingency questions (Figure 7.20) to direct respondents to questions that, given previous responses, are applicable to them.

The use of arrows and inset boxes to highlight follow-up questions is a useful way of avoiding confusion when using contingency questions. Computer-based surveys automatically take respondents or interviewers to the next applicable question (see Chapter 8).

Instructions

To help the questionnaire flow, use the following types of instructions where appropriate.

- *General instructions:* These should include an introduction to the purpose of the questionnaire, assurance of confidentiality, how the respondent was chosen, how and when to return the questionnaire (where relevant).

Were you born in Australia?
1. [ ] Yes (go to question 2)
   [ ] No

(a) In what country were you born? _____

(b) For how many years have you lived in Australia?
_____ yrs.

(c) Are you an Australian citizen?
[ ] Yes
[ ] No

Now go to Question 2

Figure 7.20 An illustration of contingency questions

- *Section introductions:* When the questionnaire can be divided into subsections provide a brief introduction to each section such as 'Finally we would like to know just a little about your background so we can see how different people feel about the topics about which you've answered questions'.
- *Question instructions:* Indicate how many responses the respondents can tick (e.g. the most appropriate, as many as apply, one only).
- *'Go to' instructions:* Ensure you make use of these when using contingency questions that require respondents to skip some questions. In electronic questionnaires these skips are automated.

Use of space

To encourage people to complete a questionnaire avoid cluttering it. The following hints may help:

- Unless you are using a booklet format print questions on one side of the page only. It is too easy for people to miss questions printed on the backs of pages. The blank backs of pages are also useful for respondents to write additional comments.
- Provide a column about 2.5 centimetres wide on the right-hand side for computer coding for paper questionnaires (see Chapter 9).
- Leave sufficient space for open-ended questions. In electronic questionnaires the space for open questions can automatically expand to accommodate any length open response.
- List alternative responses *down* rather than across the page.
- In electronic questionnaires you should consider placing just one or two questions on a screen. Even if this means you use many screens it does not make the questionnaire look 'thicker' or longer and therefore does not make the questionnaire appear more onerous.

The task of questionnaire layout has been made easier with software specially designed for producing questionnaires (Web Pointer 7.4).

Order of questions

A good questionnaire is one in which there is a logical flow to questions. The following nine points provide some guidelines.

1 Commence with questions that respondents will enjoy answering.
   a These should be easily answered questions.
   b Factual questions should be used initially.
   c Do not start with demographic questions such as age, marital status, etc.

WEB POINTER 7.4 ***Software for producing questionnaires***

The task of questionnaire layout has been made easier by the power of widely available word processors. Specialised software that has been developed for electronic surveys has made the process even simpler. This software can produce both electronic and professional looking paper questionnaires. Demonstration versions of the software are available for download from the internet.

SphinxSurvey	www.scolari.co.uk/sphinx/sphinx.htm
SurveyWriter	www.surveywriter.com/HomePage.html
SurveyTracker	www.surveytracker.com/
SurveyWin	www.raosoft.com/products/interform/index.html
Survey Said	www.surveysaid.com/
Infopoll Designer	www.infopoll.com/
More	My website www.social-research.org provides additional software links and advice for questionnaire design.

   d Ensure that the initial questions are obviously relevant to the stated purpose of the survey.
2 Go from easy to more difficult questions.
3 Go from concrete to abstract questions.
4 Open-ended questions should be kept to a minimum and, where possible, placed towards the end of the questionnaire.
5 Group questions into sections. This helps structure the questionnaire and provides a better flow.
6 Make use of filter questions to ensure that questions are relevant to respondents.
7 When using a series of positive and negative items to form a scale, mix up the positive and negative items to help avoid an acquiescent response set.
8 Electronic questionnaires can randomise the order of questions within sections for each respondent to help minimise the effect of question order within sections.
9 Where possible try to introduce a variety of question formats so that the questionnaire remains interesting.

Setting up for coding

If the data are to be analysed by computer and you are using a paper questionnaire (rather than an electronic questionnaire—see Chapter 8) it is useful to prepare for this by allocating codes to responses in the questionnaires so that a number is printed in the questionnaire next to responses. This precoding is possible only for forced-choice questions (see Figure 7.19). Depending on the way in which data are to be entered for computer analysis, computer column numbers may need to be allocated to each variable in the right-hand margins (see Chapter 9).

WEB POINTER 7.5 ***Questionnaire examples on the internet***

A valuable way to learn about questionnaire layout and the differences that follow from different modes of questionnaire administration is to look at actual examples of questionnaires. The following sites provide some examples.

Postal questionnaires	http://ssda.anu.edu.au/SSDA/CODEBOOKS/AES98/aes98cbk.rtf (the actual questionnaire is on pp. 156–81 of this file) http://ssda.anu.edu.au/SSDA/CODEBOOKS/ACRS99/d1018pub.pdf (the questionnaire is on pp. 112–32 of the file)
Face-to-face questionnaires:	
Social capital questionnaire	www.worldbank.org/poverty/scapital/library/ugquest.pdf This questionnaire was used in a face-to-face survey on social capital in Uganda.
Healthy retirement questionnaires	www.umich.edu/~hrswww/center/qnaires/download.html This site enables you to download and view a large number of questionnaire modules from the US-based longitudinal study on health and retirement. It provides many examples of questionnaire layout and instructions appropriate to a face-to-face survey.