

COGNITIVE INTERVIEWING PRACTICE

EDITED BY
DEBBIE COLLINS

NatCen
Social Research that works for society

 **SAGE**

Los Angeles | London | New Delhi
Singapore | Washington DC

FIVE

DEVELOPING INTERVIEW PROTOCOLS

JO D'ARDENNE

5.1 Introduction

The aim of this chapter is to provide practical guidance on how to design a cognitive interview protocol. Interview protocols provide the cognitive interviewer with all the information he or she needs to carry out the interview. The purpose of a protocol is to:

- a) highlight the aims of the testing;
- b) explain how test questions should be administered; and
- c) indicate which cognitive interviewing techniques should be used (e.g. think aloud and/or probing), and where appropriate, which scripted cognitive probes.

In this chapter we will discuss:

- The importance of defining measurement objectives and testing aims prior to designing the protocol
- How to decide which questions to test
- How to administer the test questions
- How to train participants to think aloud
- How to write different types of cognitive probes
- The factors to consider when deciding on which cognitive interviewing techniques to use in your test.

Example interview protocols are shown at the end of this chapter.

5.2 Define and clarify measurement objects

Before you start designing your cognitive interviewing protocol it is imperative that you are familiar with the measurement objectives of the questions you are testing. One purpose of cognitive interviewing is to test whether survey questions work in the way intended by the researcher. This means that participants *understand* the questions in the intended way and are *willing* and *able* to provide the required information. However, the above assumes that you, the researcher, have a clear idea of what the intended meaning of the question is and what information is required. If you are not clear about what the question is trying to measure how will you be able to test whether the question is meeting its measurement objectives?

It is very easy to write survey questions that initially appear straightforward but that become ambiguous on closer inspection. Let us look at the following example:

Q1. Have you visited your doctor in the last month?

- Yes
No

On the face of it this question looks relatively simple. However, on closer inspection a number of ambiguities arise.

- Should participants only include times when they visited the doctor about their own health or should they include taking their children or other dependants to see the doctor?
- Should participants include visits to the doctor's surgery, even if their appointment was not to see the doctor (for example, if they went to see the practice nurse)?
- Should participants include times a doctor has visited them, for example on a home visit?

If you are designing your own research project you need to ensure that each question has a well-defined measurement objective and that you have an appreciation of what you want participants to include and exclude when answering. To do this you may want to go back to your original research aims and objectives and reflect on how each question relates to these and to the analysis you would like to carry out once the survey data have been collected.

If you are working within a research organisation, conducting research on behalf of other people, you may need to return to the project's sponsor to clarify exactly what they see as being the measurement objectives. Be aware that the sponsor may not always have fully defined measurement aims. Therefore time should be built into the questionnaire development process to discuss the precise purpose of each question prior to testing commencing.

Including the measurement objective of each question with the test question, as shown in Figure 5.1 will help you to formulate your cognitive probes. In addition,

Q1. Have you visited your doctor in the last month?

Yes

No

[Q1 Measurement objective: To capture whether the respondent has seen a doctor, either on behalf of themselves or someone else, in the last month. This could be a GP or a consultant. Visits to other health practitioners (nurses, opticians, dentists) should be excluded.]

Figure 5.1 Example measurement objective

you will be able to revisit this document when writing up the findings from your cognitive interviews (see Chapter 9).

5.2.1 Understanding concepts within questions

Survey questions often involve the use of concepts. Sometimes you will be the question designer and therefore you will (or should) know how concepts are being defined. However, there may be occasions when you are being asked to test someone else's questions. In these circumstances it is helpful to check concept definitions. Let's look at another example.

Q2. How many portions of fruit and vegetables did you eat yesterday?

0

1–2

3–4

5+

Q3. How many units of alcohol did you drink yesterday?

_____ *units*

Questions 2 and 3 contain the concepts 'portion of fruit' and 'unit of alcohol'. These concepts have a specific meaning and in testing the questions you may want to check that participants understand them in the way the question designer intended. Knowing what the definitions are will help you design probes that explore participants' understanding of them.

5.3 Setting cognitive interviewing aims

Once you are clear about the measurement objectives of the questions you are testing, you can start to formulate your test aims. Cognitive interviewing aims are typically a list of all the areas you would like to explore in order to check that the question is

working as intended. Areas to investigate are often based on the four-stage model of survey response (Tourangeau, 1984) described in detail in Chapter 1.

Figure 5.2 shows how the four stages of Tourangeau’s model can translate into cognitive interviewing aims.

Area of investigation	Aims of cognitive interviewing
1. Comprehension	<ul style="list-style-type: none"> • To explore comprehension of key terms within the question, such as ‘doctor’. • To explore comprehension of the question as a whole, for example that respondents understand that they have to count up all the times they have been to the doctor in the past month.
2. Retrieval	<ul style="list-style-type: none"> • To establish whether respondents can recall the required information (e.g. visits to the doctor in the past month). • To establish whether respondents restrict their recall to the specified reference period (e.g. the last month).
3. Judgement	<ul style="list-style-type: none"> • To explore respondent strategies when answering. For example, do respondents try to recall each time they visited the doctor during the past month or do they just take a guess? • To explore the boundaries of what respondents include and exclude within their answers. For example, how they define ‘visit’.
4. Response	<ul style="list-style-type: none"> • To explore whether the question is considered sensitive or embarrassing and the impact this may have on the data collected. For example, could respondents edit their answers or refuse to provide an answer? • To explore whether respondents are able to map their ‘in mind’ answer onto the answer categories available. • To check whether any answer categories are missing from the list provided. • To establish whether respondents consider answer categories to be mutually exclusive (if only one category can be selected).

Figure 5.2 Using the four-stage model to formulate cognitive interviewing aims

Now let’s apply the four-stage model to the visiting a doctor question and set some aims for the testing of this question. Figure 5.3 shows the interview protocol – the test question, the measurement objective of the question and the aims for the cognitive interviewing.

The aims of testing will vary depending on the nature of the question being tested. For instance, in Figure 5.3 there is no aim related to testing answer categories. This is because ‘Yes/No’ answer categories are considered relatively straightforward and therefore testing these may not be a priority. If you were testing a question with a longer list of answer categories or a visual answer scale, setting an aim to explore response would be more important.

Not all cognitive interviewing aims are related to testing specific questions. Cognitive interviews can also explore more general factors related to the survey experience. For example, additional aims in a cognitive interview might be to:

Q1. Have you visited your doctor in the last month?

Yes

No

[Q1 Measurement objective: To capture whether the respondent has seen a doctor, either on behalf of themselves or someone else, in the last month. This could be a GP or a consultant. Visits to other health practitioners (nurses, opticians, dentists) should be excluded.]

Aims of testing Q1

- To explore **comprehension** of the phrase ‘visited the doctor’ and whether it is in line with the measurement objective.
- To explore whether respondents can **recall** if they visited a doctor in the last month.
- To explore which **judgements** respondents make when deciding what to include/exclude within their answer.
- To explore sensitivity and whether respondents are **comfortable** answering this question.

Figure 5.3 Aims of testing

- explore what factors could increase people’s willingness to take part in the survey;
- establish whether participants are able to accurately provide factual information or whether the survey would benefit from a request to see documentation (such as employer pay slips or household bills);
- explore whether participants think any important topics are missing from a questionnaire. For example, if your questionnaire aims to measure satisfaction with a service, participants may be able to think of areas they want to give feedback on (related to their satisfaction) that you have not thought to include. This can happen even if you have carried out questionnaire development work (such as depth interviews or focus groups) prior to designing your questions (see Chapter 2);
- [for self-administered questionnaires] collect opinions on the attractiveness of the questionnaire and how this could be improved;
- [for self-administered questionnaires] collect opinions on the usability, in terms of the instructions, layout and flow of the questionnaire and how this could be improved;
- [for paper questionnaires] establish whether participants are able to follow navigational instructions and, if not, how these can be made clearer.

We recommend that all interviewing aims, both those related to specific questions and more general aims, are recorded on the interview protocol. This will help keep the interview focused if spontaneous probes are used.

5.4 Deciding which questions to test

If you are testing a long questionnaire it may not be possible to test every question in a single cognitive interview. The average length of a cognitive interview

is around one hour and this includes the introductions, the administration of the survey questions and subsequent probing.

Don't be tempted to test too much in one go. Taking part in a cognitive interview requires a sustained amount of concentration from participants. If interviews last for longer than one hour participants may get fatigued. This means that the questions you test towards the end of the interview may not be explored as fully as you would like.

There is no golden rule about the maximum number of questions that can be tested in one cognitive interview. The total number of questions you can test is dependent on:

- the complexity of the questions;
- whether participants will be asked all the questions or whether some questions will be asked only of those who provide a particular answer at a previous question;
- the extent of cognitive probing envisaged; and
- whether you are hoping to explore any other issues in the interview, for example views on the visual presentation of the questions.

We have found that typically in a one-hour cognitive interview you will be able to test around 15–20 individual questions. When assessing how many questions you want to test, be aware that irrespective of question numbering, a question is one request for information. So a series of 10 agree/disagree statements should be treated as 10 questions. If your questionnaire contains more than 20 questions you will need to decide how this will be managed in testing. You can either:

- a) prioritise testing certain questions over others; or
- b) **create two (or more) different interview protocols.** For example, if you have 40 questions you wish to test, you could create one protocol to test the first 20 questions and another protocol to test the second 20 questions. Each individual interview will therefore only involve the testing of 20 questions (with participants being allocated to the first or second protocol).

If you decide you would prefer to produce just one protocol there are a number of ways you can prioritise which questions to test. Priorities include:

- **New questions that you have written from scratch** rather than those you have adapted from existing sources. That said, it is worth noting that just because a question has been used in a survey before does not mean it has been tested. Additionally, if the question has been tested with one audience it does not mean it will still be suitable if you use it with a different audience. Context can also have an impact on how participants understand questions. For example, a question which aims to measure long-standing illness, disability and health conditions is likely to give different estimates when asked on a health survey to when it is asked in a survey about employment.
- **Questions that aim to measure unusual or technical concepts.**
- **Key filter questions.** By this we mean questions that determine subsequent questions participants are asked.
- **Questions presented in an unusual or innovative format,** for example questions with pictorial answer categories.

Questions that are less of a priority for testing may include those that:

- **need to be retained in their current form for comparison purposes;** and
- **have been tried and tested in a similar context.**

But how will you know if questions have been tested before? There are a number of online resources that list both existing survey questions and question testing reports (see below).

Useful online resources

The UK Data Service: www.ukdataservice.ac.uk

The UK Data Service provides access to data collected from a range of major UK and cross-national social surveys. Questionnaires can be downloaded from most of the archived surveys. Technical reports (detailing what questionnaire pretesting was done) are provided in some instances and not others.

Q-Bank: www.cdc.gov/qbank

Q-Bank stores questions that have been tested and used primarily in US Federal surveys. A search function allows users to identify questions on different topics and all questions are linked to a question evaluation report. A list of cognitive testing reports is also available.

It is worth noting that just because questions are in the public domain does not mean they have been tested. Similarly, even if they have been tested the results of testing are not always easily accessible. Question testing findings, if published at all, are often only included as an appendix or a technical annex.

If you are working for a research organisation, conducting research on behalf of other people, you will need to discuss testing priorities with the funder. It is important to manage expectations about the amount of material that can be tested in one interview. One way to set priorities is via an expert review (see section 2.2.3).

You may need to include some additional questions in your cognitive protocol on top of the questions you wish to test. For example:

- **Demographic questions.** We often include demographic questions at the start of an interview, to ease people in to the interview and to provide contextual information. Often, these demographic questions may have already been asked as part of the screening process (see section 4.3.1). However, even if they have been asked before you may wish to include them again to check that the information you have is correct. This is particularly important if you have used a third-party recruiter (such as a gatekeeper or a recruitment agency).

- **Contextual questions.** Questions may appear to have a different meaning if they are viewed in isolation. If you want to add a couple of new questions to an existing questionnaire module or series you should consider asking participants some or all of the existing questions in that module or series. This is done to replicate the context in which the questions would be asked/read in the actual survey.

Even though these demographic questions and context questions do not need to be probed on they will still take up some time in the interview. The time it takes to administer these questions needs to be taken into account when planning your interview protocol. Figure 5.4 presents a scenario we encountered where it was important to include contextual questions.

Study background

We were asked to test some new questions designed for use in the Smoking, Drinking and Drug Use survey (SDD). SDD is an annual survey conducted in schools across England with young people aged 11–16. Respondents are asked to complete a self-completion questionnaire under 'exam hall' conditions. The new questions we were testing were on how young people obtain cigarettes, and exposure to second-hand smoke.

Designing interview protocols

- We did not want the test respondents to complete the entire SDD questionnaire booklet during the cognitive interview as this contains over 200 questions. Many of the existing questions were not related to the new questions we were testing (for example questions about alcohol and drug-use).
- However, we didn't want to test the new questions in isolation because we felt this could influence the way they were viewed by young people. For example, earlier questions had set the context of what to include as 'smoking' and had contained navigational instructions so that smokers were asked additional questions.
- Therefore we created a shorter test version of the questionnaire. This contained the original cover page, instructions and approximately 30 questions (12 of these were the new test questions and the rest were contextual questions from the existing smoking module).
- The visual presentation and formatting of the new questions was in keeping with the existing questionnaire.
- The interview protocol described how the interviewers should introduce the questionnaire. For example, interviewers were asked to explain to participants that in the real survey young people would fill in the questionnaire at school under 'exam hall' conditions, that their name is not included on the questionnaire and that no one would be able to see their answers.
- Interviewers observed participants, with their consent, completing the questionnaire.
- If consent to being observed was given, participants were asked to complete the test questionnaire whilst thinking aloud. If permission was not given participants were allowed to fill in the questionnaire privately.
- After completing the test questionnaire participants were asked additional probes about the new questions to collect further details about comprehension, perceived sensitivity and so forth.

Figure 5.4 Adding context questions and placement instructions to an interview protocol

Contextual questions mean that participants do not view the text questions in isolation but rather as a part of a larger survey experience.

5.5 Administering the test survey questions

The interview protocol should make it clear how the test questions should be administered. In a real-life survey setting questions are administered in a standardised way. By doing this researchers can be confident that variations in answers reflect real differences between participants rather than artificial differences caused by the questions being presented in different ways. In cognitive interviews it is important to replicate survey conditions in terms of the way questions are administered. How this is done will vary depending on the survey mode of questionnaire you are testing (see section 10.4). Your cognitive interview protocols will vary depending on whether the survey questions you are testing are meant to be *interviewer administered* (read out by an interviewer with supplementary written material if required) or *self-completed* (where the participant fills in a questionnaire either on paper or on a computer).

5.5.1 Interviewer-administered questions

When testing *interviewer-administered* survey questions a copy of the test questions should be included in the protocol. The test questions should be read out **exactly as worded** and in the way the questions would be asked in the actual survey. This means you need to specify in the protocol:

- If pre-specified answer options should be read out;
- Whether single or multiple answers can be coded for each question; and
- Any questionnaire navigation instructions, for example showing if additional survey questions need to be asked based on the participants' answer to a previous survey question.

Figure 5.9, at the end of this chapter, shows a protocol for testing interviewer-administered questions.

5.5.2 Self-administered questionnaires

The protocols for testing *self-administered* questionnaires (either paper questionnaires or web questionnaires) differ slightly to protocols for testing interviewer administered questions. When testing self-administered questionnaires participants should be asked to fill out the questionnaire. This can either be the whole questionnaire (containing all the survey questions, including those that are not the focus of the testing) or a test questionnaire containing only the questions you are testing and any necessary instructions/questions to set the context.

When writing protocols for testing self-administered questionnaires you need to consider *how* you will give the test questionnaire to the participant and what you will say. Chapter 10 provides more details.

5.6 Cognitive interviewing techniques

There are a number of different cognitive interviewing techniques that you can use, either on their own or in combination, to test your questions or questionnaire. In this section we look at features of these techniques that should be documented in your interview protocol.

5.6.1 Observations

When testing self-administered questionnaires you may want to observe participants whilst they complete the form. Observation is helpful in framing spontaneous (unscripted) probes to ask at a later point in the interview, without interrupting the participant as they complete the questionnaire. Examples of spontaneous observational probes include:

I noticed you skipped question seven. Can you explain why that was?

You changed your answer at question nine, why was that?

Without including a checklist it is easy to forget areas to follow-up on later in the interview. An observation checklist, such as the one shown in Figure 5.5, can be a useful tool onto which you, as the interviewer, can record observations.

It is worth noting that we include observational checklists as an aide-mémoire to help subsequent probing, rather than as a source of numerical data to be used

Observation	Details
<input type="checkbox"/> Did the respondent skip any questions?	
<input type="checkbox"/> Did the respondent change an answer at any point?	
<input type="checkbox"/> Did the respondent ever select two answer categories when only one was required?	
<input type="checkbox"/> Did the respondent ever trigger an error message? <i>[Web questionnaire only]</i>	
<input type="checkbox"/> Did the respondent ever click the 'help' button? <i>[Web questionnaire only]</i>	

Figure 5.5 Observation checklist

in analysis. Therefore, when reporting our findings we would focus on the reasons why a question was skipped rather than the number of times it was skipped.

5.7 Using think aloud

Think aloud allows you to examine the participant's own reactions, views and thought processes in real time with minimal prompting from the interviewer. There are pros and cons of the think aloud method for collecting cognitive data (see Figure 5.6).

Advantages	Disadvantages
<ul style="list-style-type: none"> • It is relatively easy to train participants to think aloud. In our experience most participants will be able to provide some information via this method. • Thinking aloud collects participant-initiated rather than interviewer-initiated data. Reports of problems initiated by participants may be more reliable compared to problems elicited by probing (see Conrad and Blair, 2009). • Think aloud data is collected at the point at which the question is answered. When data is collected retrospectively (e.g. through probing) participants may be more likely to forget or over elaborate on problems they encountered. 	<ul style="list-style-type: none"> • Not all participants feel comfortable thinking aloud. Think aloud data can vary in quality between individuals. • Thinking aloud could interfere with the task being undertaken. • Participants may not be providing an actual 'stream of consciousness' when thinking aloud. Their monologue may be edited and refined prior to verbalisation. This means problems could still be overlooked. • The think aloud data, in itself, may not address all the research questions you wish to explore.

Figure 5.6 Advantages and disadvantages of the think aloud method

When deciding on whether think aloud is an appropriate technique to use on your study, ask yourself the following questions.

- Are the questions you are testing likely to be very sensitive? Will participants feel comfortable having to think aloud their answers? In these circumstances probing may be more appropriate, to check understanding of the questions without needing to refer to the answers provided.¹
- Will the act of thinking aloud interfere substantially with how participants go about answering the questions or filling in the questionnaire? For example, we were once asked to test a number of questions designed to assess memory. The decision was made not to use think aloud for testing these questions as the act of verbalising thoughts out loud may have interfered with the task (see Example 5.1).

¹We used this approach when testing questions on sexual behaviour, which were to be included in a self-administered module in the National survey of Sexual Attitudes and Lifestyle (Natsal). Please see refer to Example 10.3. in Chapter 10 for further details.

Example 5.1 Testing memory assessment tasks: a rationale for not using think aloud in certain cases

Study background

We conducted cognitive interviews to test a battery of cognitive functioning items to be used in Understanding Society, a large, longitudinal survey following the lives of individuals in private households in the UK funded by the Economic and Social Research Council. The cognitive functioning items included tests on prospective memory and immediate and delayed word recall. Items on working memory and numerical ability were also tested.

The aim of the cognitive interviews was to ascertain whether the test instructions were clear. The cognitive functioning items were to be used in a survey context where a large number of respondents do not speak English as a first language. Therefore one aim of the cognitive interviewing was to establish whether performance in the memory tests was influenced by language ability rather than not being able to remember things per se.

Study design

- We conducted 43 interviews with people in their own homes. Of these 21 interviews were conducted with people who primarily spoke English at home and 22 interviews were conducted with people who primarily spoke a different language at home. Participants also varied in terms of their age, sex, ethnicity and self-rated English proficiency.
- In total we tested seven types of cognitive functioning task. The memory tasks tested included:
 - *A prospective memory test: prospective memory can be described as 'remembering to remember'. Prospective memory was assessed using a 'clipboard' task (see Huppert et al., 2006). Participants, at 'time one' were shown a clipboard. They were then told that at some point during the interview this clipboard would be handed to them and, when this happened, they needed to write their initials on the top right-hand corner. Later in the interview (at 'time two') participants were handed the clipboard. Interviewers were instructed to record whether the participant remembered to write their initials and, if so, whether they remembered to write them in the correct position.*
 - *A word recall test: participants had a list of words read out to them by the interviewer. Straight after this they were asked repeat back all the words they could recall. At a later point in the interview participants were asked which words from the original list they could still remember. Performance in both the immediate word recall test and word delayed recall test was recorded.*
- The decision was made not to collect think aloud data. This was because we thought that thinking aloud could actively interfere with completing the tasks. For example, on the prospective memory test we thought participants may be more likely to remember instructions if these were verbalised during the

think aloud process. For both the prospective memory and the word recall task we thought thinking aloud could increase the amount of time between 'time one' and 'time two'. This would effectively increase the task difficulty for people who were more loquacious think-alouders.

- Consideration was given to when probing should occur in order to minimise the impact of probes on task performance. For the most part probes were conducted concurrently (after each task was completed). If a task was conducted in two stages (e.g. with a 'time one' and a 'time two' stage) probing was not conducted until both stages of the task were complete.
- When designing the probes the main focus was to investigate comprehension of instructions. Other elements of the question and answer model (e.g. recall, judgement and response categories) were not explored in probing in any detail.

Study findings

- We found that performance could be influenced by language ability. For example, there were instances of people failing the prospective memory task as they did not understand the word 'initials'. Subsequently the prospective memory task was dropped from the survey and the instructions for other tasks were edited to aid comprehension.
- For more information on this study please refer to Gray et al. (2011) and Uhrig et al. (2011).

Your interview protocol should specify whether or not think aloud techniques should be used and how you would like think aloud to be introduced. There is no hard and fast rule about how the think aloud should be introduced. However, in our experience it is often necessary to tailor the way you introduce the think aloud, depending on the participant you are talking to. Here are some general pointers.

- **Keep explanations of think aloud simple.** You might wish to say one or more of the following.

'When you answer the questions I would like you to tell me what you are thinking...'

'Please say 'out loud' anything that comes into your head whilst you are answering...'

'Please tell me what is going through your mind as you work out each of your answers...'

- **Provide a simple practice example.** We sometimes use the 'windows example' described by Gordon Willis (2005), which involves the participant counting up how many windows they have in their home. This example is useful as it can be applied to nearly everyone you are likely to meet: your participants are likely to live somewhere with windows but are unlikely to know 'off-the-top-of-their-head' how many windows there are. An example of how to introduce the think aloud using the windows example is shown in Figure 5.10 at the end of this chapter.
- **Use praise and positive feedback** to encourage reticent participants to continue talking during the practice.

In some cases participants will be able to think aloud perfectly naturally without much prompting. In other cases you may have to demonstrate the think aloud example by doing it yourself for the participant, to indicate what you mean.

It is possible to conduct cognitive interviews where participants are instructed to think aloud in some parts of the interview and not others. For example, we have designed protocols that include a large number of contextual questions that are not the focus of the testing. In these instances we have successfully instructed participants to only think aloud at certain points (e.g. the test questions and not the contextual questions). If you plan to use think aloud in some parts of your interview then let participants know this at the outset. Your protocol should clearly indicate when think aloud should and should not be used.

5.8 Probing

Think aloud, by itself, may not provide you with all the information you need to meet your test aims because:

- not all participants may be comfortable thinking aloud and therefore may not provide sufficient information; and
- even participants who are naturally good at thinking aloud may not verbalise all of their thought processes.

In this section we provide examples of different types of probes you may wish to include. We will discuss using scripted versus spontaneous probes, general probes and probes that are theory-driven. A summary of best practice when probing is provided in section 5.8.3.

5.8.1 Scripted vs. spontaneous probes

When designing your interview protocol you will have to decide whether you are going to include scripted probes or whether you will rely solely on spontaneous probes.

- **Scripted probes** are designed pre-interview and are included in the cognitive interviewing protocol.
- **Spontaneous probes** are not included in the protocols. They are generated *during* the interview. Spontaneous probes still need to address the research aims but there is more flexibility in how they are framed and they are more responsive to specific areas mentioned by participants.

One advantage of using scripted probes is that you can ensure they conform to the best practice principles discussed in section 5.8.3 (e.g. they are phrased in a neutral manner, that they are simple and so on). It is far more difficult to ensure that these best practice principles are being adhered to if you are required to compose probes 'on the hoof'. In addition, it will be easier for you to ensure that you address all your testing aims if you have a list of scripted probes to follow.

Another advantage of using scripted probes is that you can ensure that the same areas are being explored with every participant in a consistent way. This will make the data management phase of the analysis much more straightforward (see Chapter 7) and can be helpful when carrying out cognitive interviews in cross-national, cross-cultural and multilingual settings (see Chapter 11).

However, despite this we believe that to get the most useful information cognitive interviewers should have freedom to generate their own (spontaneous) probes during the interview. This allows you to:

- a) explore relevant issues that arise, even if these issues have not been predicted in the protocols; and
- b) seek clarification on what the participant has already said if the meaning is initially unclear.

In addition, we believe that, as a cognitive interviewer, you should be able to tailor the probes you ask dependant on the participant. Participants will vary in terms of their language skills and their ability to express themselves verbally. You should tailor the language you use when probing as appropriate in order to facilitate communication.

Therefore we recommend that the probing part of the interview should be *semi-structured* and involve both scripted probes and opportunities for spontaneous probing, in line with what Beatty (2004) and Beatty and Willis (2007) recommend. If multiple people are involved in conducting the interviews, all interviewers should be trained in how to generate good-quality probes (that are open, balanced and clear – see section 5.8.3). Likewise, all interviewers should always be briefed on the aims of the question testing to ensure that any spontaneous probes collect relevant and useful information. The use of unscripted probes is not without its risks, and their use needs careful consideration when undertaking cognitive interviews in different languages (see Chapter 11).

5.8.2 Respondent-driven vs. theory-driven probes

Probes can vary in terms of their specificity.

Respondent-driven probes are general probes that aim to gather participants' initial reactions to a question. They are called respondent-driven probes because they allow participants to comment on any aspect of the question they chose: Examples of respondent-driven probes might be:

How did you find answering these questions?

Were they easy or difficult? Why do you say that?

The advantage of using respondent-driven probes is that you can gain an insight into how questions function from a participant's point of view, without any

potential bias caused by the interviewer. In this sense data collected from general probes is similar to findings generated by think aloud. Respondent-driven probes are non-directive as participants can discuss any issues they wish, including unanticipated problems.

The disadvantage of respondent-driven probes is that they may not yield full information about how a question is performing. This can occur when participants are not motivated to provide detailed responses to general probes. In addition general probes can result in problems going **undetected**. For example, after a general probe a participant may describe how he or she found a question easy to answer and provide some form of justification of why this was the case. However, on further probing it may transpire that the same participant has misunderstood a key word and not realised it.

Theory-driven probes seek to investigate specific problems that might have occurred based on the researcher's pre-existing knowledge about the measurement aims of the questions.

Figure 5.7 illustrates several examples of theory-driven probes based on the four-stage mode of survey response (Tourangeau, 1984).

It is worth reminding ourselves that Tourangeau's four-stage model is not the only theoretical framework that can be used to explore the question and answer process. For example, Willis (2005) has developed a framework for writing probes based on his Questionnaire Appraisal System. Under this framework probes can be formulated to explore:

- a) Instructions
- b) Clarity
- c) Assumptions
- d) Knowledge/ Memory
- e) Sensitivity/ Bias
- f) Response categories
- g) Other issues.

Whichever model you choose the advantage of using theory-driven probes is that all potential areas of concern can be explored in a systematic way. By looking at the question prior to testing you can start to predict what types of problem it might have and then draft probes to establish whether or not these problems arise in practice.

The disadvantage of using theory-driven probes is that they may lead to artificial or 'artefact findings'. Artefact findings occur when participants start to think about questions in a new way as a result of the cognitive interviewing. Directive probing can influence participants' perceptions of a question (Conrad and Blair, 2009). Let us return to the example used earlier in this chapter of visiting your doctor as an illustration, shown in Example 5.2.

Area of investigation	Aims of cognitive interviewing	Example probes
1. Comprehension	<ul style="list-style-type: none"> To explore comprehension of key terms within the question. To explore comprehension of the question as a whole, for example to see whether comprehension is influenced by the questions' length or sentence structure. 	<ul style="list-style-type: none"> What did you understand by the term 'X' when answering this question? What does the term 'X' mean to you? In your own words what do you think this question is trying to ask? [This type of probe is known as 'Paraphrasing']
2. Retrieval	<ul style="list-style-type: none"> To establish whether respondents can recall the required information. To establish whether respondents restrict their recall to the reference period specified in the questions. For example; In the last seven days how many units of alcohol have you drunk?. 	<ul style="list-style-type: none"> How easy or difficult was it to remember 'X'? What time period were you thinking about when answering? From when until when? When did you last do 'X'?
3. Judgement	<ul style="list-style-type: none"> To explore respondent strategies when answering. For example, do respondents use an accurate calculation strategy or just take a guess? To explore whether the question is considered sensitive or embarrassing and the impact this may have on the data collected. For example, do respondents edit their answers or refuse to provide an answer? 	<ul style="list-style-type: none"> How did you work out your answer to this question? How accurate would you say your answer is? How did you feel about being asked this question? How comfortable would you feel answering this question in a telephone survey about 'X'? Why?
4. Response	<ul style="list-style-type: none"> To explore whether respondents are able to map their 'in mind' answer onto the answer categories available. To check whether any answer categories are missing from the list provided. 	<ul style="list-style-type: none"> How easy or difficult was it to select an answer from the options provided? Why? Why did you tick that particular box (choose that particular answer), rather than one of the others? Are there any categories missing from the options provided or do they cover everything? What is missing?

Figure 5.7 Examples of cognitive probes based on the four-stage model of survey response

Example 5.2 Probes and artefact findings

Q1. Have you visited your doctor in the last month?

Yes

No

[Q1 Objective: To capture whether the participant has seen a doctor, either on behalf of themselves or someone else, in the last month. This could be a GP or a consultant. Visits to other health practitioners (nurses, opticians, dentists) should be excluded.]

Aims of testing Q1:

- To explore **comprehension** of the phrase ‘visited the doctor’ and whether it is in line with measurement objective.

Suggested probe:

What did you understand by the word ‘doctor’?

In this example it might be that a participant only thinks about their GP when the survey question is first asked (in line with measurement objectives). However, on hearing the probe ‘*what did you understand by the word “doctor”?*’ the participant might start to think of a broader range of things compared to when he or she first answered the question. For example, the participant might start considering allied health professionals (who should be excluded) as a result of reflecting on the probe. For this reason, it is important that probes are framed to discover what the participant was thinking about *when answering the question*. An improvement of the above probe would therefore be:

Suggested probe:

When answering this question what did you understand by the word ‘doctor’?

In practice we typically use a combination of respondent-driven and theory-driven probes in cognitive interviewing protocols. Respondent-driven probes should come before more specific theory-driven probes, as this allows participants to articulate their initial views on a question without any steer from the interviewer and improves error detection (Conrad and Blair, 2009). Theory-driven probes can then be used to investigate additional areas that have not been previously discussed.

5.8.3 Writing cognitive probes

Cognitive probes need to collect information that is unbiased, clear and that addresses the research aims. This section describes a number of best practice principles to adhere to when writing cognitive probes.

Firstly, a good probe (like a good survey question) should be neutrally phrased. Biased probes can increase the likelihood of problems going undetected. For example, if a probe reads ‘*Was that easy?*’ or ‘*How easy was it to remember X?*’ participants may be naturally inclined to respond in the affirmative, regardless of how they found the question. Similarly, probes can encourage participants to report problems that may not have occurred, for example in the following wording: ‘*You seemed puzzled, was it difficult?*’ Even if it seems clear to you that the participant was experiencing difficulties you should not make assumptions. This could lead you to detecting problems that do not occur in practice or exaggerating problems that do occur. Probes should be either:

- **Neutral**, where no particular mental state should be inferred in the question, e.g. ‘*How did you find that?*’; or
- **Balanced**, where any states inferred should be balanced with reference to an opposite state, e.g. ‘*Did you find it easy or difficult?*’

Secondly, the ideal scenario for cognitive interviewing is that participants should be doing roughly 80% of the talking, with the interviewer (you) doing the remaining 20%. To ensure participants do most of the talking you should use open probes that require more than a single word answer. Following up a closed answer to a probe with ‘*Why?*’ or ‘*Why do you say that?*’ is a straightforward way to encourage participants to give more expansive responses.

Thirdly, probes should be kept relatively simple. Probes that are long-winded or double-barrelled (where a single probe asks about multiple things) or have multiple clauses are likely to be misunderstood, in the same way a complicated survey question can be misunderstood. Again, if probes are misunderstood there is a danger that the cognitive interviewing could create artificial findings in that participants describe confusion as a result of the probes, rather than the question you are testing.

Finally, when writing probes you should always refer back to the aims of the cognitive testing. By doing this you can check that all your aims are being directly addressed in the probes you write, and that there are no superfluous probes collecting information irrelevant to your aims.

A summary of best practice when writing probes is shown in Figure 5.8.

Good probing	Bad probing
<ul style="list-style-type: none"> ✓ Neutral/balanced ✓ Open ✓ Simple ✓ All testing aims are covered ✓ Relevant to testing aims 	<ul style="list-style-type: none"> ✗ Leading ✗ Closed ✗ Complicated or double-barrelled ✗ Some testing aims are overlooked ✗ Irrelevant to testing aims

Figure 5.8 Summary of good (and not so good) probing techniques

5.8.4 When to probe

As well as considering what probes to include, some thought needs to be given to the placement of the probes.

Concurrent probing is when probes are administered directly *after each survey question*. The advantage of concurrent probing is that participants should be able to recall what they were thinking about when answering the question. The disadvantage of probing after each question is that it breaks up the natural flow of the survey questions. When testing self-administered documents it is difficult to carry out concurrent probing without continually interrupting the participant.

Retrospective probing is when probes are administered *once all the survey questions have been answered*. The advantage of using this type of probing is that the probes do not interfere with the flow of the survey questions. This type of probing means that the experience of answering the questions more closely mirrors the context of an actual survey. The disadvantage of retrospective probing is that participants may not be able to remember what they were thinking about when they answered each question. This can lead to participants post-rationalising their survey answers, leading them to describe things that were not actually thought about when they first answered.

The decision as to whether to use concurrent probing, retrospective probing or a combination of both will vary depending on what you are testing. The following factors should be considered when deciding on when to probe.

- a) **Questionnaire structure.** If the questionnaire you are testing has distinct sub-sections on different topics you may wish to probe after each sub-section rather than at the end of the questionnaire.
- b) **Question similarity.** If you are testing a number of questions containing the same word or phrase comprehension probes should only be asked after all the questions containing the repeated phrase have been administered. For example, if you are testing two questions using the phrase 'general health', probing on understanding of 'general health' should occur after both questions have been administered, otherwise there is a danger that probing could influence understanding of the second question.
- c) **Questionnaire mode.** This is discussed in section 10.4.5.

As discussed previously, we want to try and minimise the occurrence of artefact findings when probing. When reviewing your protocol you should consider whether the placement of your probes could influence how participants react to the questionnaire. If probes could influence the understanding of later questions or the navigation of the questionnaire their positioning should be reconsidered.

5.9 Using different cognitive interviewing techniques together

If a decision has been made to use think aloud and probing within the same interview, you will need to consider how your interview protocol will be arranged in terms of which technique the interviewer should use first. It may be that you wish to do either, or a combination, of the following.

- a) Think aloud and probing for individual questions.
- b) Think aloud for certain questions but probing for others.
- c) Think aloud exclusively for certain sections.
- d) Only probing for certain sections.

When using think aloud and probing to explore a question (1), think aloud should be encouraged first (before any probes) and *during* the question administration (before or whilst the survey answer is being provided). This is because, as already noted, think aloud is a respondent-driven technique and information elicited via it should not be influenced by the interviewer in any way. Once the participant has finished thinking aloud, probes can be administered. This arrangement of the techniques should be clearly explained to participants at the start of the interview.

If interviewers are to only use think aloud or only probing (2) for a question or during a section of the interview protocol (3 or 4), then this should be explained to participants.

5.10 Other information to include in the protocol

Your interview protocol should provide some information on how to introduce the interview to participants. For example, protocols should set out what the interviewer should say when introducing the study at the start of the interview and in obtaining participants' informed consent to take part. More details on how to introduce a cognitive interviewing study are given in section 6.5.

Similarly, your protocol should also summarise any extra steps you need to carry out at the end of each interview. For example, protocols should contain reminders to give participants incentives (if these are being used), check if participants have any questions and direct participants to further sources of information (see section 6.8).

5.11 Testing the protocols

It is recommended that you conduct a ‘test’ cognitive interview before you finalise your protocol. These test interviews can happen early on, in advance of the main fieldwork, with volunteers from within your organisation or university, or even with a colleague or a friend. A lot will be learned through doing some early interviews and they will help you to revise the probes, write the instructions and pre-empt what to do about unexpected issues that may arise.

A test interview will allow you to check the length of the interview. In general, we recommend that the maximum length of a cognitive interview should be about one hour. After an hour participants may begin to tire, and the information they give you will be less useful. If your test interview shows that your protocols are too long you have two options.

- a) **Drop some of the material from the protocol.** You may need to drop some questions from the test or not probe on particular questions. This will involve you revisiting your testing priorities and reviewing your testing aims.
- b) **Develop two (or more) versions of the protocol** and test different versions on different participants (see section 5.4).

Testing the protocol will also enable you to see whether there are any missing or unclear instructions or poorly phrased probes. Getting a sense of how the interview flows is very helpful at this stage and will help you to fine tune the interview protocol and develop appropriate guidance for other cognitive interviewers working on the project.

5.12 Chapter summary

Your cognitive interviewing protocols should:

- highlight the measurement objectives and the question-testing aims;
- explain how the test questions should be administered; and
- suggest which cognitive interviewing techniques to use and when.

We recommend using a combination of all or some of the following in your protocols: observations, think aloud and respondent-driven and theory-driven probes. The four-stage model of survey response (Tourangeau, 1984) or alternative question and answer models can be used as a starting point when writing theory-driven probes.

When writing probes you need to make sure they are neutral, open, simple and relevant to your testing objectives. Scripted probes can be supplemented

with spontaneous probes to clarify what participants have said and to explore issues that are not predicted by the protocol.

5.13 Examples of cognitive interview protocols

Some examples of cognitive interview protocols are shown in this section.

- Figure 5.9 shows how you might train the participant to think aloud.
- Figure 5.10 shows an example protocol from testing interviewer-administered questions.

Think aloud training

- First, train the respondent to think aloud.
- *You may use the windows example or another example of your choice.* An example script is shown below if this is required.
- For this project we would like respondents to think aloud for **all** of the test questions

Example script

- Let me explain a little bit more about how the interview will work. I'm going to read ask/give you some questions. When you hear them, I want you to tell me whatever comes into your mind. This is called 'thinking out loud'.
- We've found that it helps to have some practice at doing this. So let me give you an example. Let's say I was asked a question about 'how many windows are there in my house'.
- If I was thinking out loud, I would say... *[GIVE EXAMPLE] :..well, there's one window in the kitchen ...and then in the living room, there are two windows. Well I say windows, they are panes of glass – they don't open. Not sure if I should include those. I think I will. And then ... (etc., etc.)'*
- Now, let me ask you the same question. Think about how many windows there are in your house. As you count up the windows, tell me what you are seeing and thinking about.

Further prompts if required:

- **Please say, out loud, what you are thinking....**
- **What's going through your mind?**

Note: In some cases you may want to design protocols that only use think aloud at certain points.

Figure 5.9 Protocol for think aloud

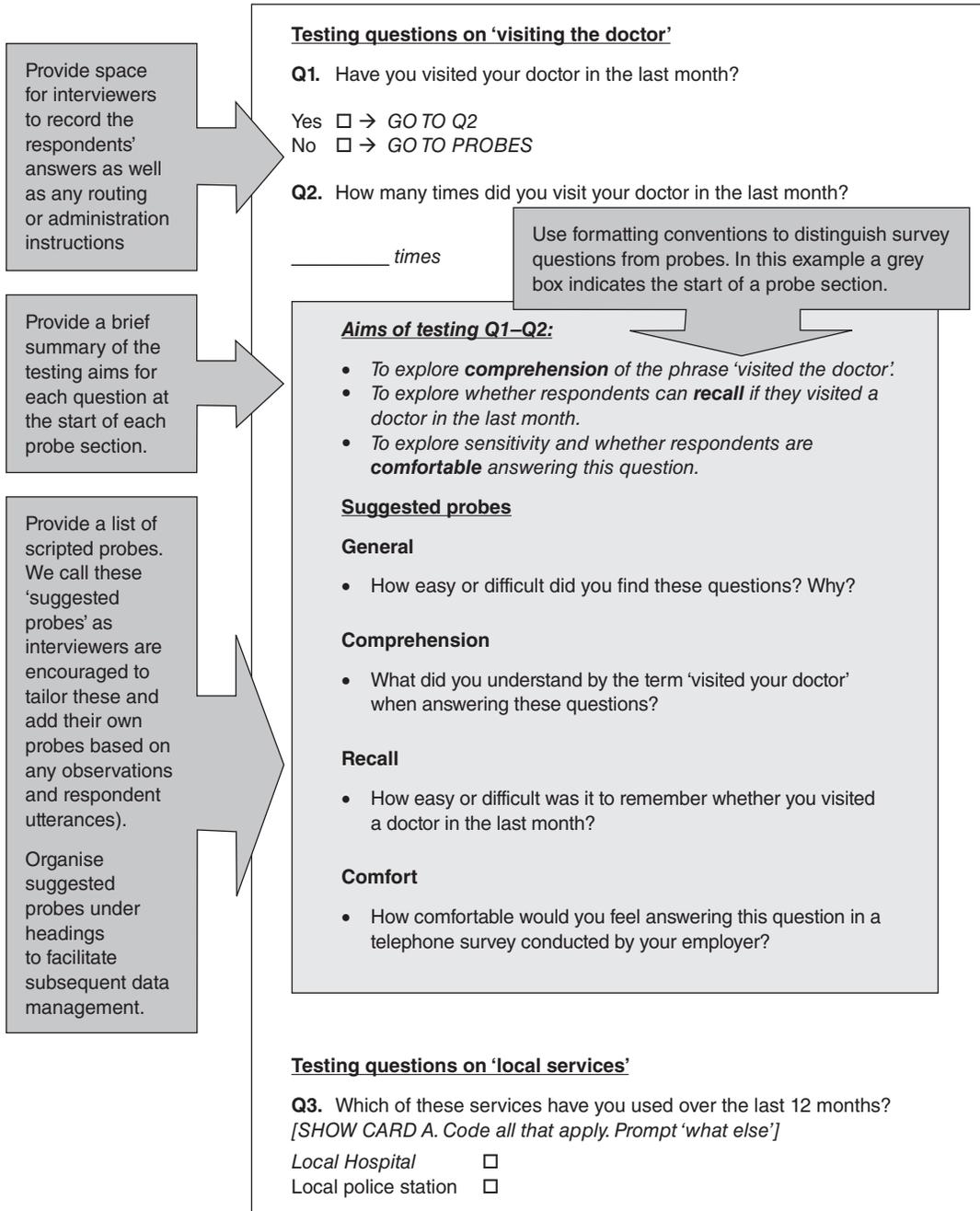


Figure 5.10 Example cognitive protocol for interviewer-administered questions

Your protocols should also provide details on how you want the interview to be introduced and closed. We recommend that you test your protocol before it is finalised.

References

- Beatty, P. (2004) 'The dynamics of cognitive interviewing', in S. Presser, J.M Rothgeb, M.P. Couper, J.T. Lessler, E. Martin, J. Martin and E. Singer (eds), *Methods for Testing and Evaluating Survey Questionnaires*. Hoboken, NJ: John Wiley & Sons. pp. 45–66.
- Beatty, P. and Willis, G. (2007) 'Research synthesis: the practice of cognitive interviewing', *Public Opinion Quarterly*, 71: 287–311.
- Conrad, F. and Blair, J. (2009) 'Sources of error in cognitive interviews', *Public Opinion Quarterly*, 73(1): 32–55.
- Gray, M., d'Ardenne, J., Balarajan, J. and Uhrig N. (2011) *Cognitive Testing of Wave 3 Understanding Society Questions*. Understanding Society Working Paper Series, No.2011– 03 available at: <https://www.understandingsociety.ac.uk/research/publications/working-paper/understanding-society/2011-03.pdf>
- Huppert, F.A., Gardener, E. and McWilliams, B. (2006) 'Cognitive function', in J. Banks, E. Breeze, C. Lessof and J. Nazroo (eds), *Retirement, Health and Relationships of the Older Person in England: The English Longitudinal Study on Ageing (Wave 2)*. London: Institute for Fiscal Studies. pp. 217–30.
- Tourangeau, R. (1984) 'Cognitive science and survey methods: a cognitive perspective', in T. Jabine, M. Straf, J. Tanur and R. Tourangeau (eds), *Cognitive Aspects of Survey Design: Building a Bridge Between the Disciplines*. Washington, DC: National Academy Press. pp. 73–100.
- Uhrig, N., Gray, M., McFall, S., d'Ardenne, J. and Balarajan, M. (2011) *Say What? Testing of Cognitive Functioning Tests with Non-Native English Speakers in the UKHLS*. Paper presented at 4th European Survey Research Association conference, Lausanne, Switzerland.
- Willis, G. (2005) *Cognitive Interviewing: A Tool for Improving Questionnaire Design*. Thousand Oak, CA; Sage.