gesis Leibniz-Institut
für Sozialwissenschaften

GESIS Survey Guidelines

Cognitive Pretesting

Timo Lenzner, Cornelia Neuert, & Wanda Otto

Please note that this Gesis Survey Guideline is an old version. A revised version can be found on the Gesis Survey Guidelines website.

December 2016, Version 2.0

Abstract

In survey research, cognitive pretesting is generally considered an essential prerequisite for successful questionnaire development, and thus for the quality of the survey data. This contribution provides an introduction to questionnaire pretesting and focuses especially on two key aspects – namely, the planning and implementation of cognitive pretests. The following questions are addressed, in particular: What is a pretest, and why should questionnaire pretests be conducted? What pretesting methods are available, and which one should you choose? What techniques are used in cognitive pretests? How are cognitive pretests conducted? How are cognitive pretests analysed?

Citation

1. What is a pretest, and why should questionnaire pretests be conducted?

The term *questionnaire pretest* broadly refers to the evaluation or testing of a questionnaire before it is administered in the actual survey. Pretests are thus an essential part of the questionnaire design process. The purpose of a pretest is to provide information about:

- The comprehensibility of the questions

    Does the meaning that respondents associate with a question correspond to the meaning intended by the researcher? Do respondents interpret the meaning of a question in the same way?

- Difficulties that respondents have with their task

    How difficult is it for respondents to understand and answer the question? Is the subject matter of the question perhaps unfamiliar or sensitive?

- Respondent interest in, and attention to, individual questions

    Do fatigue effects manifest themselves during the interview/while the questionnaire is being completed? Do respondents think that (individual) questions are redundant?

- Frequency distributions of the responses

    Is the full range of the scale used?

- Context effects and problems with the question order

    Do earlier questions influence responses to subsequent questions?

- Interviewers' problems

    Can interviewers recognise clearly what they are supposed to read out and what they are not supposed to read out?

- Technical problems with the questionnaire and with interview aids (e.g., lists, showcards)

- The duration of the interview/questionnaire completion

(see Converse & Presser, 1986; Porst, 2000).

All these aspects provide important information about whether the questionnaire functions as it should, and thus about the quality of the data collected with it. To ensure "good" data, desk-based appraisal of the questionnaire is generally not enough because, as Sudman and Bradburn (1982: 283) pointed out: "Even after years of experience, no expert can write a perfect questionnaire." Only with the help of empirical pretests is it possible to check whether survey questions actually measure what they are supposed to measure and whether they yield reliable and valid responses.

2. What pretesting methods are available, and which method should you choose?

A very diverse range of pretesting methods are available for evaluating questionnaires. They include, for example, conventional pretesting, cognitive interviewing, behaviour coding, respondent debriefing,

group discussion, expert review, eye tracking, and web probing (Behr, Braun, Kaczmirek, & Bandilla, 2013; Campanelli 2008; Faulbaum, Prüfer, & Rexroth, 2009; Prüfer & Rexroth, 1996). Each of these methods has specific strengths and weaknesses. The decision in favour of one or more of them depends essentially on the field of application and the knowledge interest. In what follows, we focus on two pretesting methods: (1) conventional pretesting, which should always be applied before any survey, and (2) cognitive pretesting, which should, ideally, be applied as part of the questionnaire design process, in addition to, and before, conventional pretesting.

2.1   Conventional pretesting

In a conventional pretest, the questionnaire is administered under conditions that simulate as realistically as possible those that will prevail in the main survey (rehearsal piloting). In other words, a conventional pretest is a simulation of the main survey, as it were, with a sample of between 10 and 200 respondents (Prüfer & Rexroth, 1996). Conventional pretests should be conducted in the mode (face-to-face, telephone, postal, online, etc.) that will be used in the main survey. To avoid influencing the response process, the respondents are not usually informed that the survey is a pretest.

The objective of a conventional pretest is to check the practicability of the interview process as a whole and the functionality of the entire questionnaire. As a rule, this yields reliable information about (1) technical defects in the questionnaire (e.g., defective filter questions or problems that interviewers have administering the questionnaire), (2) frequency distributions of the responses, and (3) the average duration of the interview/questionnaire completion. Moreover, a conventional pretest can also yield limited information about the way respondents understand the questions. For example, when piloting interviewer-administered questionnaires, the interviewers could be instructed to make a note of, and to report, difficulties and striking behaviour on the part of the respondents during debriefings. In general, however, conventional pretesting is a *passive method* in which the interviewer or the pretester merely observes the respondent without actively probing his or her understanding of the questions. As a rule, the fact that respondents have difficulties understanding a certain question is discovered only if they point this out themselves. Hence, a conventional pretest usually yields only little, unsystematic, and superficial information about the respondents' understanding of the questions.

2.2   Cognitive pretesting

Cognitive pretesting (or cognitive interviewing) belongs to the *active pretesting methods* because the participants' approach to answering the questions is actively probed and investigated (Beatty & Willis, 2007; Prüfer & Rexroth, 2005; Willis, 2005). Cognitive pretesting is usually conducted during the questionnaire design phase in order to obtain an insight into the cognitive processes that take place when questions are being answered:

- How do participants interpret questions or terms?

- How do they retrieve information and events from memory?

- How do they arrive at a decision as to how they should respond?

- How do they assign their "internally" determined responses to formal response categories?

The actual objective of cognitive pretesting is to obtain information about a wide range of questionnaire problems. Cognitive pretests are particularly suitable for testing the comprehensibility of questions, identifying problems that respondents have answering the questionnaire, establishing the causes of these problems, and generating suggestions for improvement on the basis of these findings.

2

In contrast to conventional pretesting, the focus in cognitive pretesting is on testing individual questions rather than evaluating the questionnaire as a whole. Cognitive pretests are generally suitable for all survey modes. In other words, they can be applied irrespective of whether the subsequent survey is to be administered face-to-face, by telephone, by paper and pencil, or online. The main survey benefits most when the questionnaire also undergoes conventional pretesting after the cognitive pretest in order to test its functionality under realistic conditions that simulate those of the main survey.

3. What techniques are applied in cognitive pretests?

3.1   The think-aloud technique

The think-aloud technique involves asking participants to "think aloud" and to verbalise all the thought processes that lead, or led, to their response. The aim is to reveal the response process, and thus also any problems that participants may have understanding a question. The think-aloud method can be used either while the question is being answered (concurrent think-aloud) or after it has been answered (retrospective think-aloud). In the case of concurrent think-aloud, the interviewer introduces the question to be tested with an instruction such as the following (Porst, 2014): "While you are answering the following question, can you tell me what you are thinking, or what is going through your mind? Please also mention things that may appear to you to be unimportant. The question is: ...."

3.2   Probing techniques

Probing is a technique that involves asking participants one or more follow-up questions (probes) about terms, questions, or responses. This enables additional information to be gained about the way in which participants understand the questions. Depending on the knowledge interest, different types of probes can be applied: (1) comprehension probes (e.g., "What do you understand by 'a job with a high level of responsibility' in this question?"), (2) category-selection probes (e.g., "Can you explain why you chose this answer?"), (3) information retrieval probes (e.g., "How did you remember that you had gone to the doctor [...] times in the past 12 months?"), and (4) general/elaborative probes (e.g., "Can you explain your answer in more detail?"). Probes can be administered concurrently (after the participant has answered the survey question) or retrospectively (after he or she has answered the whole questionnaire).

3.3   Paraphrasing

Paraphrasing involves asking participants to repeat the question in their own words after they have answered it ("Can you repeat the question I just asked you in your own words?"). Ideally, this verbalisation yields information about whether, or how, the participant understood the question and whether this understanding corresponds to that of the researcher. It should be noted, however, that the fact that someone does not do a good job of repeating the question in their own words does not necessarily mean that they have not understood it. Generally, paraphrasing is not suitable for short factual questions because participants are usually able to remember them word-for-word, and one therefore obtains only information about the participants' memory performance rather than their understanding of the question.

3.4   Confidence rating

This technique involves asking participants to assess the reliability of their response after they have answered the question and to report how confident they are that they have given the correct answer ("How sure are you that you went to the doctor [...] times in the past 12 months?"). If participants report that they are not quite sure whether their response is correct, they should always be asked about the reason for this uncertainty.

3.5   Sorting

Sorting techniques are used to investigate the way in which participants assign terms or situations to certain categories (e.g., which descriptions of an accident situation do participants classify as a *traffic accident*, and which descriptions do they not classify as such). We distinguish between free and dimensional sorting:

- Free sorting involves asking participants to group specific items according to their own criteria.

- In the case of dimensional sorting, participants are asked to sort items according to predefined criteria.

This technique presents researchers with the challenge of constructing in advance categories that adequately cover all substantive aspects of the term in question.

There are no universal rules as to how varied the cognitive techniques applied in a pretest should be. On the one hand, a certain amount of variety is definitely a good way to avoid tiring the participants and directing their attention away from the survey questions and towards the anticipated probes. On the other hand, merely for the sake of variety one should not apply techniques that may not yield any new knowledge. Rather, the application of a cognitive technique should always be determined by the knowledge interest of the researcher or by the behaviour of the participant.

4. How are cognitive pretests conducted?

There are no firm rules as to how cognitive interviews should be conducted. Ideally, they should take place in a quiet, closed room and should be recorded with a video camera (or, if this is not possible, with a dictaphone). The advantage of a video recording is that the interviews can be analysed not only acoustically but also with regard to striking visual phenomena (facial expressions, gestures). However, recording the interviews (whether it be audio or video) should not be dispensed with, as it facilitates and improves both the implementation and the analysis of the interviews.

There are different practices and recommendations with regard to the number of interviews that should be conducted in the context of a cognitive pretest. As a rule, between five and 30 interviews are conducted per pretest (round) because the most serious question problems can usually be identified on the basis of a relatively small number of interviews (Willis, 2005). On the other hand, Blair and Conrad (2011) demonstrated that conducting more cognitive interviews than are typically carried out increases the probability of uncovering further significant question problems. However, when one considers the large volume of verbal text data produced in the context of cognitive interviews, and the fact that these data must be subjected to analysis, conducting more than 30 interviews per pretest appears quite impracticable. Because of such cost-benefit considerations, researchers tend therefore to work with around 20 participants per cognitive pretest. If sufficient resources are available, it is recommended that the questions that have been revised on the basis of the findings of the cognitive interviews

should be tested in a further round. This iterative approach enables the effectiveness of the revisions to be evaluated (Prüfer & Rexroth, 2005; Willis, 2005).

The degree of standardisation of cognitive interviews ranges from (almost) completely unstructured to (almost) completely standardised. When an unstructured approach is adopted, the interviewer applies the cognitive techniques completely spontaneously and only as a reaction to the behaviour of the participant. In the case of a completely standardised approach, the cognitive techniques to be used are scripted in advance in an interview protocol, and the interviewers are instructed to administer the protocol in the most standardised way possible. We recommend a mixture of the standardised and the non-standardised approaches. An interview protocol should be drawn up in advance that contains the questions to be tested, the cognitive techniques to be used in each case, and space for the interviewer's notes and comments. If necessary, further unprepared cognitive techniques and (probing) questions can be spontaneously added during the interview. It should be left up to the interviewers to continue probing until they are sure that they have obtained all the necessary information.

Cognitive interviews should be scheduled to last between 60 and a maximum of 90 minutes. If they last any longer, the concentration and motivation of the participants (and the interviewers) will wane significantly. Depending on the number of cognitive techniques applied, between 20 and 25 questions, or items, can be tested during this time. The pretest participants should generally have the same characteristics as the respondents in the main survey (in terms of age, sex, education, etc.). Typically, a quota sample is selected for cognitive interviewing. It is not necessary to draw a random sample because the main objective of cognitive interviewing is to uncover problems with the questions rather than provide as precise an estimate as possible of the frequency with which these problems occur in the population.

5. How are cognitive pretests analysed?

Before analysing cognitive pretests, it is useful to transcribe the individual recordings and to generate a case-specific list of all participant utterances. Such a list should include the following information: (1) the responses to the tested questions, (2) spontaneous participant utterances regarding the question, (3) participant responses and reactions to cognitive probes, and (4) remarks by the cognitive interviewer.

Several different methods of evaluating cognitive interviews are available. The simplest and fastest method is the **informal analysis** of the data. Here, the analyst decides in the case of each participant utterance whether or not it indicates that a problem exists with the question. However, this approach bears the risk that question problems will be very subjectively assessed and the pretest findings will lack verifiability.

The **formal analysis** of the data can be carried out quantitatively or qualitatively. The **quantitative analysis** of cognitive interviews involves the use of coding schemes (e.g., DeMaio & Landreth, 2004). Here, codes are assigned to behaviours (or verbal utterances) of the interviewers and the participants, for example: (1) participant has difficulty understanding the question, (2) participant does not understand certain words, (3) participants understand the question differently, and (4) participants have difficulty recalling the subject matter addressed in the question. Hence, the quantitative analysis of interviews is a systematic and objective method that allows inter-rater reliability to be assessed and a quantitative analysis of the data to be conducted (e.g., counting the frequency of different problem types). The decisive disadvantages of this method are the risk of losing information by reducing text data to codes and the fact that suggestions for improvement cannot be made on the basis of the codes (rather, one must go back to the verbal data).

The **qualitative analysis** of cognitive interviews involves the application of qualitative procedures such as the constant comparative method (CCM). This method is particularly suitable for explorative data analysis and hypothesis generation. It comprises three work steps:

1. Open coding: The verbal data of the participants are openly coded by topic, and initial categories are created.

2. Axial coding: The researcher endeavours to integrate the categories created from the verbal data and checks whether there are group differences in the assignment of the categories.

3. Selective coding: In the last step, superordinate topics that connect the categories are specified. When doing so, a hypothesis or theory is formulated that describes the phenomena that a survey question captures.

A good presentation of the application of this method in the context of the analysis of cognitive interviews can be found in Ridolfo and Schoua-Glusberg (2011).

6. References

Behr, D., Braun, M., Kaczmirek, L., & Bandilla W. (2013). Testing the validity of gender ideology items by implementing probing questions in web surveys. *Field Methods, 25,* 124-141.

Beatty, P. C., & Willis, G. B. (2007). Research synthesis: The practice of cognitive interviewing. *Public Opinion Quarterly, 71,* 287-311.

Blair, J., & Conrad, F. G. (2011). Sample size for cognitive interview pretesting. *Public Opinion Quarterly, 75,* 636-658.

Campanelli, P. (2008). Testing survey questions. In E. de Leeuw, J. Hox, & D. Dillman (Eds.), *International Handbook of Survey Methodology* (pp. 176-200). New York: Erlbaum.

Converse, J. M., & Presser, S. (1986). *Survey questions: Handcrafting the standardized questionnaire.* Beverly Hills, CA: Sage.

DeMaio, T., & Landreth, A. (2004). Do different cognitive interview methods produce different results? In S. Presser, M. P. Couper, J. T. Lessler, E. Martin, J. Martin, J. M. Rothgeb, & E. Singer (Eds.), *Methods for testing and evaluating survey questions.* (pp. 88-108). New York: Wiley.

Faulbaum, F., Prüfer, P., & Rexroth, M. (2009). *Was ist eine gute Frage? Die systematische Evaluation der Fragenqualität.* Wiesbaden: VS Verlag.

Porst, R. (2000). *Praxis der Umfrageforschung.* Stuttgart: Teubner.

Porst, R. (2014). *Fragebogen. Ein Arbeitsbuch.* Wiesbaden: VS Verlag für Sozialwissenschaften.

Prüfer, P., & Rexroth, M. (1996). Verfahren zur Evaluation von Survey-Fragen: Ein Überblick. *ZUMA-Nachrichten, 39,* 95-116.

Prüfer, P., & Rexroth, M. (2005). Kognitive Interviews. *ZUMA How-to-Reihe, No. 15.*

Ridolfo, H., & Schoua-Glusberg, A. (2011). Analyzing cognitive interview data using the constant comparative method of analysis to understand cross-cultural patterns in survey data. *Field Methods, 23,* 420-438.

Sudman, S., & Bradburn, N. M. (1982). *Asking questions: A practical guide to questionnaire design.* San Francisco, CA: Jossey-Bass.

Willis, G. B. (2005). *Cognitive interviewing: A tool for improving questionnaire design.* Thousand Oaks, CA: Sage.