Medical Sociology Unit, Hannover Medical School

Testing questionnaires with cognitive methods: part I, Probing questions

For a long time preparing questionnaires for use in surveys has been more an art than an evidence-based endeavour. Expert interviews and standard pretests (interviews performed under field conditions with minimal interventions at the side of interviewers) had been the methods of choice (Converse & Presser 1986), but now they are considered as insufficient if applied alone (Tourangeau et al. 2000). They allow to estimate the time frame for performing an interview, whether the forms can be handled, and whether a study design is feasible. However, standard pretests do not permit to uncover ambiguities unless subjects are asking for clarification or unless they are giving inappropriate responses.

In the 1980s research on methods was combined with traditions from cognitive psychology in order to improve questionnaire design (Sudman et al. 1996; Tourangeau et al. 2000). The main topics are whether respondents understand survey items how they should and whether they are answering accordingly (Dillmann 2000). In the literature two procedures are discussed as promising ways: *probings* and *thinkalouds.* Additionally, supplementary methods like sorting tasks or focus groups can be performed for obtaining further clarification (Geyer 2003).

Probings are performed in order to examine the appropriate understanding of survey questions. After having answered, subjects are requested to explain their responses (Kurz et al. 1999, 23: 62–82). With *comprehension probings* subjects' understanding of the expressions used is explored. This may lead to the conclusion that even seemingly simple questions may carry ambiguities: If patients' satisfaction with medical treatment in outpatient settings is examined, the following question can be asked: *"How often do you have been to the doctor within the last two months?"* The understanding of what is meant by "doctor" may vary considerably and lead to different conclusions. Some subjects may think of a doctor as a person, or they might have been in a medical practice without actually having seen the physician, or they may have been in a hospital and seen the doctor there.

The question *"How satisfied are you with the German health care system?"* poses even more ambiguities since "health care system" is a broad and ambiguous expression requiring clarification. Some subjects may understand it as a question for the health insurance system, others may have the way of financing medical care in mind, others may judge the quality of health care delivery of in- and outpatient treatment.

With *category selection probings* subjects are asked to give reasons for their responses on rating scales: In doing so, they have to make explicit what meaning response categories are carrying for them, they have to reproduce what pieces of information had been used and how they have been combined and weighted: "*Please explain why you have put your tick in the box* [repeat the scale position chosen]?".

Information retrieval probings are directed towards the process of how responses are generated: *"While you have answered this question, what did you think until you have given your answer?"*

Paraphrasing of questions: Respondents are repeating questions with their own words. This permits to check whether they were understood in the same way as intended by the researchers. *"Please repeat the question that I have just read with your own words. What was the question?"* If paraphrases are very different from the original question, it can be concluded that it may not have been understood. The same holds if respondents stay close to original formulations, if they are repeated word by word, or if they changed only one expression.

Confidence ratings (DeMaio & Rothgeb 1996: 177–196) can be performed as a supplementary procedure in order to examine the certainty ratings are done with. (Example: *"Please rate the certainty of your judgement on the following scale."*) This may be applied to survey interviews where respondents give ratings on scales, if interviewer-based ratings are performed, the confidence of the background information can be assessed.

Soz.- Präventivmed. 48 (2003) 133–134
0303-8408/03/020133–02
DOI 10.1007/s00038-003-3007-5

All these methods have to be employed in personal interviews in order to permit flexible reactions and more detailed exploration.

In practice, probings should be performed after the question program had been compiled, after all questions have been constructed and before the field work is about to start. However, probings may even be used in ongoing studies or after the analyses of data for obtaining further clarification after survey questions have led to ambiguous results. All newly constructed questions should be tested. It is also recommended to test questions that are about to be applied in samples that are drawn from populations where they have not yet been used. The analyses of probings can involve a qualitative analysis of responses, but also quantitative accounts of how often misunderstandings came up or how often certain responses were given in order to decide whether questions have to be modified.

After a series of probings and subsequent modifications of questions it may become necessary to run a second trial because some changes may not turn out to be real improvements (Kurz et al. 1999). Thus working on questionnaires is an iterative process and it is up to the researchers to decide when to stop.

Testing questionnaires with cognitive methods have not yet been established as standard procedures, thus no precise recommendations on sample sizes permitting a good balance between work and profit can be given. In a recent test of a questionnaire designed for use in the general population 17 interviews had been performed (Kurz et al. 1999), Mittag et al. (2003) used a case number of 20, in earlier studies 36 (Jobe et al. 1996) or up to 100 probing interviews (Groves et al. 1992) were completed. As a general rule one should keep in mind that in surveys to conduct with general population samples large variations in how questions are interpreted have to be expected. Consequently, pretest studies will have to cover the whole spectrum of subjects and ways of interpretations that can potentially be encountered in the population. A case number of 17 may not be sufficient, and 36 may be appropriate. In more homogeneous samples 10 to 20 may already be enough to reveal the scope of ways of how questions can be understood or misunderstood.

With the procedures outlined promising ways of testing and improving survey questions have been described. The development of such techniques is still at the beginning, but encouraging results on their applicability have been produced.

References

Converse JM, Presser S (1986). Survey questions: handcrafting the standardized questionnaire. Newbury Park: Sage.

DeMaio TJ, Rothgeb JM (1996). Cognitive interviewing techniques: in the lab and in the field. In: Schwarz N, Sudman S, eds. Answering questions. San Francisco: Jossey Bass: 177–96.

Dillmann DA (2000). Mail and internet surveys: the tailored design method. New York: J. Wiley.

Geyer S (2003). Empirische Forschungsmethoden in den Gesundheitswissenschaften. Weinheim: Juventa-Verlag.

Groves RM, Fultz NH, Martin E (1992). Direct questioning about comprehension in a survey setting. In: Tanur JM, ed. Questions about questions. New York: Russell Sage Foundation: 49–64.

Jobe JB, Keller DM, Smith AF (1996). Cognitive techniques in interviewing older people. In: Schwarz N, Sudman S, eds. Answering questions. San Francisco: Jossey-Bass: 197–219.

Kurz K, Prüfer P, Rexroth M (1999). Zur Validität von Fragen in standardisierten Erhebungen. Ergebnisse des Einsatzes eines kognitiven Pretestinterviews. ZUMA-Nachrichten *23*(44): 62–82.

Mittag O, Böhmer S, Deck R, et al. (2003). Fragen über Fragen: cognitive survey in der Fragebogenentwicklung. Soz Praventiv Med *48*: 55–64.

Sudman S, Bradburn NM, Schwarz N (1996). Thinking about answers. San Francisco: Jossey Bass.

Tourangeau R, Rips LJ, Rasinski K (2000). The psychology of survey response. Cambridge: Cambridge University Press.

Address for correspondence

Prof. Dr. Siegfried Geyer
Medizinische Soziologie OE 5443
Medizinische Hochschule Hannover
D-30625 Hannover

e-mail: geyer.siegfried@mh-hannover.de

Soz.- Präventivmed. 48 (2003) 133–134