VU Research Portal

The effect of the question topic on interviewer behavior; an interaction analysis of control activities of interviewers

van der Zouwen, J.; Smit, J.H.; Draisma, A.R.

published in

Quality and Quantity
2010

DOI (link to publisher)

[10.1007/s11135-008-9181-z](https://doi.org/10.1007/s11135-008-9181-z)

document version

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

van der Zouwen, J., Smit, J. H., & Draisma, A. R. (2010). The effect of the question topic on interviewer behavior; an interaction analysis of control activities of interviewers. *Quality and Quantity*, 44, 71-85.
<https://doi.org/10.1007/s11135-008-9181-z>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

The effect of the question topic on interviewer behavior; an interaction analysis of control activities of interviewers

Johannes van der Zouwen · Johannes H. Smit ·
Stasja Draisma

Published online: 20 June 2008
© Springer Science+Business Media B.V. 2008

Abstract In a standardized personal interview, elderly (65+) Dutch respondents ($N = 233$), were asked detailed retrospective questions about six physical activities like walking, cycling and their performance of household tasks. Surprisingly, the proportion of inadequate answers was small, suggesting that the interviewers—four professional research nurses—have done their very best to eventually obtain adequate answers. They used three different types of control of the interview process. Firstly, an optimal execution of the prescribed ‘open loop control’, that is, precisely following the text of the questionnaire. Secondly, via the ‘feedback loop control’ of repair: if nevertheless inadequate answers are given, further probing is performed until eventually an adequate answer is obtained. Thirdly, by decomposing a general question into its components, making the question easier to answer (partial questioning). Interaction analysis of transcripts of the interviews showed that the type of control exercised by the interviewers, strongly differed by topic of the question. It appeared that question topics requiring more complex cognitive activities of the respondents, not only lead to more (need for) repair, but also to larger differences between the interviewers concerning their readiness to use partial questioning as an alternative for the open loop control as designed by the researcher.

Keywords Question topic · Retrospective questions · Survey interviews · Interviewer related measurement error · Interviewer behavior · Behavior coding · Interaction analysis

1 The interviewer: both process facilitator and source of measurement error

If, for a substantial part of the respondents, the questions in a questionnaire are hard to answer, self-administered questionnaires might be replaced by telephone or face-to-face interviews.

J. van der Zouwen
Faculty of Social Sciences, Vrije Universiteit, Amsterdam, The Netherlands

J. H. Smit (✉) · S. Draisma
Department of Psychiatry, VU University Medical Centre, A.J. Ernststraat 887,
1081 HL Amsterdam, The Netherlands
e-mail: jh.smit@vumc.nl

The interviewer is assumed to read out the questions, give explanations if requested, and note the answers of the respondent. If the initial response of the respondent is inadequate, the interviewer is instructed to make an attempt at *repair* by further probing. Sometimes the interviewer also *anticipates* problems particular respondents will encounter, by making the questions, hopefully, less hard to answer for them. In this way, the interviewer facilitates the data collection process.

However, the employment of interviewers in a survey may also form an extra source of survey measurement error (Groves 1989, Chap. 8), negatively affecting the quality of the data obtained. In this paper we focus on two types of interviewer behavior that have substantial positive and negative effects on data quality, i.e., *repair* and *anticipation* (van der Zouwen and Smit 2005, 2006). Under what conditions will these interviewer behaviors occur during an interview? More specifically, we will investigate what the effect is of the topic of the question, on the occurrence of these two interviewer behaviors.

Repair and anticipation by interviewers may be viewed as separate forms of control behavior of interviewers. Therefore, in the next section various forms of control behavior in standardized survey interviews will be discussed. These control behaviors have a clear impact on the length of the interaction sequences in interviews. This means that from the length of the interaction sequence, i.e., the number of turns taken by interviewer and respondent, the type of control can be inferred. This procedure is illustrated with data from a study on physical activities of elderly respondents.

We will investigate in this paper to what degree the question topic, in this case the type of physical activity reported, affects the use of additional control activities of interviewers.

2 Different forms of control in standardized interviews

A standardized interview can be viewed as a controlled communication process involving three actors (van der Zouwen and Smit 2005). The first one is the *researcher*, who constructs the questionnaire, and under whose responsibility the interviewers are instructed, and the collected responses are analyzed. The second one is the *interviewer*, who reads out the questions from the questionnaire, and evaluates the initial answers of the respondent. If an initial answer is inadequate, e.g., not complete, the interviewer has to probe further until she¹ obtains an adequate answer. The third actor is the *respondent*, who by means of an internal cognitive process, interprets the question, searches for relevant information, combines and evaluates this information and reports his answer (Tourangeau 1984; Tourangeau and Rasinski 1988; Tourangeau et al. 2000).

The researcher aims at getting comparable, relevant and unbiased information about the respondents. To reach this goal he has to control the interview process. In cybernetics, the science of control and communication (Wiener 1948/1965), different forms of control are distinguished (Rapoport 1982). The first form of control is *open loop control*, as exercised by the researcher. Ideally, in a standardized interview only open loop control takes place. Then the interview process directly proceeds through the following five phases: the question as worded by the researcher (phase 1) is posed about exactly so by the interviewer (phase 2), and leads to the cognitive process within the head of the respondent (phase 3), resulting in an adequate answer (phase 4), correctly noted by the interviewer (phase 5), and entered into the data file of the researcher.

¹ For the sake of readability, interviewers are indicated in this paper by she and her, whereas researchers and respondents are indicated with he and him.

But sometimes things go wrong, for example during the cognitive process within the respondent, leading to an inadequate answer. In that case, the interviewer is instructed to make an attempt at repair, for example by asking a probing question. By doing so, there emerges a *feedback loop* from phase 4 back to phase 2: the respondent is asked to reconsider and redo his initial answer (Moore and Maynard 2002). Here *closed loop control* is at work.

Previous research (van der Zouwen and Smit 2005) has shown that sometimes the interviewers ‘override’ the open loop control of the researcher. If they expect that the question as worded in the questionnaire will lead for this (type of) respondent to an inadequate answer, they may rephrase the question such that the question becomes easier to answer, hopefully leading to an adequate initial answer (Houtkoop-Steenstra 1995). By this anticipatory behavior of the interviewer, based on expectations (Hyman et al. 1954/1975) about the respondent, the ‘official’ question is replaced by one ‘invented’ by the interviewer. This alternative control can be described as *feedforward control*: on the basis of the expected inadequate answer of the respondent to the ‘official’ question by the researcher, the interviewer constructs a question that hopefully avoids the necessity of repair.

3 The impact of the form of control on the length of the interaction sequence

These three different forms of control of the interview process lead to different interactions between the interviewer and the respondent, resulting in different strings or sequences of verbal utterances of both actors. If open loop control is successfully exercised, the resulting sequence can be labeled as ‘*paradigmatic*’, that is, the “ideal sequence from the standpoint of survey design” (Schaeffer and Maynard 1996, p. 66). It “. . . consists of the interviewer posing the question as worded in the questionnaire, followed by an adequate answer of the respondent. The interviewer may close the sequence with repeating the answer or saying something like okay” (Dijkstra 2005, p. 612; Dijkstra and Ongena 2006, p. 984). In a paradigmatic sequence the number of turns is very small: two or three.

If the initial answer of the respondent is inadequate and the interviewer does an attempt at repair, the question–answer sequence becomes at least two turns longer. So the exercise of feedback control, or repair, shows itself by a longer question sequence.

Feedforward control by the interviewer, or anticipation, has different forms. A very common form is adding to the question itself a kind of candidate answer, making it a suggestive question. If the candidate answer is about correct, the resulting sequence will be as long as the paradigmatic one. Another form of anticipation is that the interviewer decomposes the topic or subject of the ‘official’ question into several components, and asks the ‘official’ question about each of these components. Examples of this *partial questioning strategy*, ‘invented’ by the interviewers are reported in van der Zouwen and Smit (2005, 2006). This form of anticipatory interviewer behavior will lead to long question–answer sequences, because for every component of the question topic a separate question must be asked and answered.

From this short exposition it appears that the type of control exercised within the interview is reflected by the length of the resulting question answer sequence: successful open loop control leads to a short paradigmatic sequence, closed loop control in the form of repair leads to some additional turns, and feedforward control, when using a partial questioning strategy, leads to many additional turns, and thus to a much longer sequence.

Since the main subject of this paper is repair and anticipation—by partial questioning—of interviewers, we focus the analysis on sequences that are significantly longer than the paradigmatic ones. The method of analysis employed will be illustrated by data from a survey after physical activities of elderly people.

4 The LAPAQ questionnaire as an illustration

We chose this material because it was shown by Schwarz et al. (1998) that elderly people often have difficulties with answering survey questions. Moreover, the specific part of the interview we selected for the analysis, mainly consisted of rather detailed retrospective questions about the frequency and duration of quite ‘mundane’ activities like walking or housekeeping, that is questions that are rather hard to answer (Knaüper et al. 1997, 2004). Nevertheless, the proportion of ‘don’t know answers’ and refusals was remarkably low. It seems that the four research nurses who interviewed these elderly respondents did all they could to get eventually adequate answers to these hard to answer questions. Thus we expected that much repair and partial questioning had taken place here.

The subject of the survey, from which part of the data is used, is “the prevention of fall accidents in older persons” (Stel 2003). It has been shown in the literature that the performance of physical activities is an important factor in the reduction of falls and fractures (Graafmans et al. 1998). For that reason, a couple of instruments is used in this survey to measure these activities (Stel et al. 2004).

One of these instruments is called the LAPAQ, a questionnaire for face-to-face interviews that covers the frequency and duration of walking outside, bicycling, gardening, light and heavy household activities, and sport activities during the previous 2 weeks.

The survey was performed within the framework of LASA, the Longitudinal Aging Study Amsterdam (Deeg and Westendorp-de Serière 1994). One year after the third data collection of LASA in 1998/1999, four trained research nurses visited 430 participants, all aged 65 years or more, at home. During this visit, they administered among other things the LAPAQ questionnaire.

Concerning each topic (abbreviated with X) of the LAPAQ questionnaire, a few questions are asked. The first question is the *base question* “Do you do X?” If the answer is positive, then for three of the six topics the base question is followed by “In the past two weeks, did you do X?” If the answer to this question is positive (too) then the *frequency question* is asked: “How many times did you engage in X during the last two weeks?” This series of questions about this topic is closed by the *duration question* “How long, on average, were you engaged in activity X?” (see the Appendix for the exact wording of the questions).² If the answer to the *base question* is ‘no’, the interviewer will skip to the next topic of the questionnaire.

The whole series of subsequent questions, responses and remarks related to one of the six topics is called a *macro-sequence*. A macro-sequence contains the question–answer sequences of the separate questions for each topic; see Box 1 for an example.³

If a macro-sequence contains all questions and answers about the topic, it is called a *complete* macro-sequence; otherwise it is called *incomplete*. For example if the respondent answers with ‘no’ to the base question (“Do you do X?”) and the interviewer goes to the next topic, the result is an incomplete macro-sequence, since the follow up questions and answers do not occur in the sequence.

² Dependent on the topic some additional questions are asked. For example, for the topic ‘gardening’ a question reads “Did you turn over the soil in your garden?”

³ Each sequence is identified by the number of the interview, followed by the number of the topic. For example, macro-sequence #11622-1 is the sequence about ‘walking’ in interview 11622. Each turn of a macro-sequence is identified by a number. The code following this number consists of two characters. The first one indicates the actor (I = interviewer; R = respondent). The meaning of the second code will be explained in the next section.

Box 1. An example of a coded macro-sequence (#11826-1)

01. Ib: I now want to ask you some questions about your physical movements pattern. Do you walk now and then? And by 'walking' we mean walking outside, to visit someone or to go shopping.^a

02. Rb: No, I only walk for my pleasure.

03. Ib: Yes, but you never go out shopping?

04. Rb: Yes.

05. Ib: So that happens now and then. And if you visit somebody; then you walk to that person? Your sister-in-law lives nearby, and then you go there on foot?

06. Rb: Yes.

07. Ip: Have you walked during the past two weeks?

08. Rp: No, not much, because it was quite windy.

09. Ip: Yes, but you have done it?

10. Rp: Yes.

11. If: And how many times did you walk during the past two weeks?

12. Rp: No.

13. Ip: And not to your brother?

14. Rp: Yes, but that is here.

15. Ip: But then you walk through the corridors?

16. Ri: Yes and he comes along too.

17. Ip: Yes, but how often have you visited him the last two weeks?

18. Rf: Two or three times.

19. If: The last 14 days; how often have you been there; two times, three times?

20. Rf: I come there once every day.

21. If: Everyday, also Saturday and Sunday?

22. Rf: Yes.

23. If: That is 14 times then.

24. Id: And how long do you walk then? If you walk through the corridors how long is that; a few minutes or is it 10 minutes?

25. Rd: I walk a quarter of an hour back and forth.

26. Id: I mean in the house, because the house is so big. Because if you go to your brother you walk through the corridors and these corridors are quite long.

27. Ri: Yes, but I do not go up the stairway.

28. Ii: No, then you go with the elevator.

29. Id: But you walk daily about 15 minutes?

30. Rd: Yes, that's right.

^a The question, as read by the interviewer, slightly differs from the one printed in the questionnaire. In the questionnaire it is directly made clear that only walking *outside* is under discussion. In the present macro-sequence, in turns 15–30 also walking inside the building, i.e., to respondent's brother is discussed and counted as 'walking'

5 The coding scheme

To investigate the topic dependency of task related interviewer behavior we need interviews with transcribed macro-sequences for all six topics. This requirement is fulfilled in 233 of the 430 interviews, so the database for the quantitative analysis consists of $6 \times 233 = 1,398$ macro-sequences, of which 931 are complete and 467 incomplete.

Each macro-sequence is divided into a series of successive turns, that is speech acts of the interviewer (I), the respondent (R); and possibly a third person (T), but only if he or she interferes with the answering process. The turns are coded with a coding scheme, using the SEQUENCE VIEWER software, developed by [Dijkstra \(1999, 2002\)](#). The coding scheme consists of only two variables (for an overview of coding schemes used for behavior coding of interviews, see [Ongena and Dijkstra 2006](#)). The first variable regards the ACTOR to whom the turn can be attributed (I, R or T). The second variable concerns the QUESTION from the questionnaire that is under discussion in this turn.

Box 2. Codes of the coding variable QUESTION

b(base): Do you do X? (Q1a, Q2a, Q3a, Q4a, Q5a, Q6a).

p(past): Did you do X during the past two weeks? (Q1b, Q2b, Q3b).

f(frequency): How many times did you do X during the past two weeks? (Q1c, Q2c, Q3c, Q4c, Q5c, Q6c).

d(duration): How long did you usually do X each time? (Q1d, Q2d, Q3d, Q4d, Q5d, Q6d).

h(household): Questions and answers related to a particular task in the household (partial questioning strategy).

m(mixed): Two questions or topics are mixed into one turn.

i(irrelevant): Questions or answers that are irrelevant to the topic.

In Box 2 the coding scheme for the variable QUESTION is presented. The question numbers Q1a, Q2a etc. refer to the LAPAQ questionnaire presented in the Appendix.

6 the number of additional turns of a macro-sequence

6.1 Correcting for the different number of questions to be asked

As mentioned in Sect. 3, the length of a question–answer sequence is an indicator of the type of control exercised by the interviewer. However, the length of a macro-sequence not only depends on task related behavior of interviewer and respondent, but also on the number of questions from the questionnaire that have to be asked and answered. So to make valid comparisons between different topics regarding the kind of control exercised, the length of the macro-sequence has to be corrected for differences regarding the number of questions that have to be asked.

As each question requires at least two turns to be asked and answered, we can subtract from the length of each macro-sequence the number of turns it would contain if the macro-sequence only consisted of paradigmatic question–answer sequences. The resulting number is called the variable ‘additional turns’ abbreviated to AT. This variable is the dependent variable in our analysis.

6.2 The distribution of the variable additional turns (AT) by topic

In Table 1 the distribution of the score on the variable AT, for each of the six topics of the LAPAQ questionnaire, is presented for all 233 interviews. About 5% of the macro-sequences have a negative AT-score, meaning that the observed sequence is even shorter than the paradigmatic one. This especially occurs with the topic ‘gardening’; for example, as the interviewer rephrases question Q3a “Do you have a garden?” for respondents living in a flat as “I see that you don’t have a garden”.

Twenty-one percent of all macro-sequences has an AT-score of zero; they contain no additional turns. This percentage is highest (41%) for the topic ‘cycling’. And another 21% has quite a small AT-score, with only one to three additional turns. So about half (47%) of the macro-sequences have so few additional turns that it is unlikely that repair, let alone partial questioning, has taken place. Probably the main form of control here is the open loop control by the researcher, as implemented by the interviewer.⁴

⁴ This does not automatically mean that these macro-sequences are all paradigmatic: much unwanted suggestive questioning can have taken place in these interviews, but that remains unnoticed here because suggestive questioning generally does not lead to additional turns.

Table 1 Distribution of the number of additional turns of all macro-sequences (AT-score) for each of the six topics of the LAPAQ questionnaire

AT score	Walking	Cycling	Gardening	Sport	Light HW	Heavy HW	Total
<0	2%	6%	15%	2%	4%	3%	5%
0	3%	41%	34%	19%	9%	21%	21%
1–3	11%	15%	34%	24%	21%	24%	21%
4–7	23%	17%	14%	35%	24%	19%	22%
8–11	20%	9%	2%	12%	17%	15%	12%
12–15	17%	4%	1%	3%	8%	6%	6%
16–19	9%	2%	0%	3%	1%	6%	4%
20–23	6%	3%	0%	1%	3%	2%	3%
24–27	3%	1%	0%	1%	4%	2%	2%
28–31	3%	1%	0%	0%	5%	2%	2%
>31	3%	1%	0%	0%	3%	1%	1%
Total	100%	100%	100%	100%	100%	100%	100%
N	233	233	233	233	233	233	1,398

$$\chi^2 = 447.7; \text{df} = 50; p < 0.001$$

The other half of the macro-sequences have AT-scores of four turns or more and therefore are certainly not paradigmatic: it is quite likely that some repair or partial questioning has taken place. The proportion of these longer sequences is very unevenly distributed over the six topics of the questionnaire: it varies from 17% (gardening) to 84% (walking).

An AT-score of 16 means that the macro-sequence is at least twice as long as the paradigmatic one. Such really long sequences did not show up with the topic gardening. On the other hand, in 16%, respectively 25%, of the sequences concerning ‘light household work’ and ‘walking’, such high AT-scores were obtained, indicating that especially these topics require a lot of additional control activities of the interviewers.

7 Content analysis of long macro-sequences

7.1 Causes of additional turns

As explained above, repair and the use of a partial questioning strategy by the interviewer, will have led to longer question–answer sequences, and thus also to longer macro-sequences. In order to study repair and partial questioning in more detail, we therefore selected for each of the six topics of the LAPAQ questionnaire the 10 longest macro-sequences. These 60 long macro-sequences contained many more turns of interviewer and respondent than the corresponding set of paradigmatic sequences. A content analysis of these macro-sequences showed the following six main ‘causes’ of these additional turns:

1. *Definition problems*: A discussion takes place between respondent and interviewer about the definition and the delineation of the physical activity mentioned in the questionnaire (for example turns 02–05 in Box 1). These problems especially occur with the topics ‘walking’ and ‘sport’. Neither from the question, nor from its explanation in the questionnaire, it is quite clear what activities fall under the label ‘walking’; e.g., is taking out the dog, or shopping, to be viewed as ‘walking’ or not? Similar problems arise with the topic ‘sport’: is walking a sport, or a means of going from one place to another?
2. *Partial questioning and probing*: The interviewer—deviating from the questionnaire—rephrases the questions and decomposes the activity into individual tasks within house-

hold work, or into different destinations for cycling or walking (e.g., turns 05 and 13). Partial questioning especially occurs with questions about activities that are hard to answer in its entirety. For ‘gardening’ and ‘sport’ there seems to be no need for using the partial questioning strategy.

3. *Completion probing*. If the interviewer uses a partial questioning strategy, it is important that all relevant household tasks, or all relevant walking or cycling destinations are dealt with. This explains the frequent use of probing questions like: “And are there other tasks you do?” Much of this probing occurs in a leading manner: “And to the doctor, you also walk?” (e.g., turn 13).
4. *Confirmation and check*. Some answers of respondents are unclear, or—at least to the interviewer—not very probable. Therefore, some interviewers are inclined to check the initial answer for clarity and validity (e.g., turn 23). This interviewer behavior is quite independent of the topic of the question.
5. *Conferring* by interviewer and respondent about what is, given the particular situation of this respondent, the best fitting answer to the question. For example, in how far has time spent with ‘doing the laundry’ to be counted, if most of the work is done by the automatic washer? Or, on what type of sport does the respondent spend more time, and thus has to be questioned first? (see also turns 24–29).
6. *A complex situation* of the respondent. The frequency questions and the duration questions assume a kind of regularity in the performance of the activities. However, sometimes the behavior of the respondent shows so much irregularity, at least over the last 2 weeks, that it is difficult to gain a clear view of his physical activities, which leads to additional turns to clarify the situation. This especially occurs with the topics ‘cycling’ and ‘household tasks’.

From Table 2 it appears that some ‘causes’ of additional turns in long macro-sequences are very unevenly distributed over the six topics of the LAPAQ questionnaire, indicating that the control behavior (especially partial questioning and the related completion probing) of interviewers will be quite topic specific.

7.2 Cognitive problems of respondents

It might be argued that the problems that occur in these long macro sequences are expressions of cognitive problems that respondents encounter during question answering. Answer-

Table 2 The different ‘causes’ of extra turns in long macro-sequences, for each of the six topics^a

	Walking	Cycling	Gardening	Sport	Light HW	Heavy HW	Total
1. Definition problems	2	0	1	3	0	0	6
2. Partial questioning and probing	8	8	0	0	10	10	36
3. Completion probing	6	6	0	1	0	1	14
4. Confirmation and check	2	2	1	2	2	1	10
5. Conferring between I and R	3	4	4	8	2	1	22
6. Complex situation of R	2	3	0	1	4	3	13
Total number of ‘causes’	23	23	6	15	18	16	101
Number of analyzed macro-sequences	10	10	10	10	10	10	60

^a In this table and the following ones, bold printed figures indicate above average numbers or proportions

ing questions comprises a series of at least four cognitive tasks according to a number of authors (Cannel et al. 1981; Tourangeau 1984; Tourangeau and Rasinski 1988; Tourangeau et al. 2000).

A first task is a proper *comprehension* of the question, which implies a unequivocal interpretation of the question topic. Especially with the topics sport and walking a lot of negotiation takes place between interviewer and respondent with respect to the interpretation of the question topic (as an example, see turns 02 until 05 in Box 1).

Complex personal situations may have an impact on the difficulty of the second cognitive task: *retrieval of information* from autobiographical memory. Especially with temporal frequency questions, the burden on memory and estimation processes may be heavy (Tourangeau et al. 2000). In case events occur irregularly in respondents' life, the use of estimation strategies becomes necessary (Groves et al. 2004, p. 218). A simple recall and count strategy, in which the number of occurrences of the intended events are remembered specifically and totaled up may be sufficient for non frequent events, for instance with respect to specific sporting appointments during a period of 2 weeks. However, a rate based strategy—in which the rate at which an event typically occurs—can be more appropriate for relatively frequent events, e.g. cooking a meal, in case the respondent still cooks regularly. The transcribed sequences show overt verbal instances of such recall and count strategies. Finally, the respondent may reside on a simple impression-based estimation, in which a vague impression is translated into a number. This strategy is supposed to be the most error prone. Knaüper et al. (2004) demonstrated that the use of such a strategy is especially likely to occur with respect to questions concerning mundane behavior of older respondents. As most topics in the LAPAQ questionnaire refer to mundane behavior of the elderly, the use of estimation strategies is likely to occur quite often.

A third cognitive task consists of the *combination* of information retrieved from memory, or a judgment about its accuracy and completeness with respect to the question. It can be argued that the conferring of interviewer and respondent about what is, given the particular situation of this respondent, the best fitting answer to the question refers to this task. Some respondents express uncertainty about the accuracy of their candidate answer verbally, and as such cause an interaction with the interviewer, in which a respondent may be probed to come up with additional information before a final answer can be constructed. Moreover, the verbal completion and check activities of the interviewer denote the fact that the interviewer interferes (or assists) in this judgment phase, since she may not be convinced about the validity of respondents' answer.

A final cognitive process involves the *mapping* of the retrieved information onto the response scale of the answer. The transcripts also demonstrate that interviewers quite often verbally assist at this phase, e.g. at turns 23 and 29 in Box 1. The transcripts show that interviewers regularly calculate a final answer overtly, and offer them as a check for the respondent, before they record the final answer.

In this sense, the expectations of the interviewer about the respondent can be understood as a mirror of the cognitive activities of the respondent, in which the interviewer tries to reflect the inconsistencies and inaccuracies in the obtained responses. These mirror activities form the control loop that guides the respondent to a recordable answer.

8 Incomplete and complete macro-sequences

In the preceding section it appeared that the questions about the six different topics were causing quite diverse cognitive problems for the respondents, and led to different forms of

Table 3 Distribution of the incomplete and complete macro-sequences over the six question topics

Topic	Incomplete macro-sequences		Complete macro-sequences		All	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
Walking	31	13	202	87	233	100
Cycling	126	54	107	46	233	100
Gardening	160	69	73	31	233	100
Sport	93	40	140	60	233	100
Light HW ^a	5	2	228	98	233	100
Heavy HW	52	22	181	67	233	100
Total	467	33	931	67	1,398	100

$$\chi^2 = 336.9; \text{df} = 5; p < 0.001$$

^a HW in this table and in the following ones is an abbreviation of ‘household work’

respondent and interviewer behaviors. In the next three sections we will complement the ‘qualitative’ content analysis of the macro-sequences, with a more quantitative analysis: in how far do different topics lead to different types of control of the interview process?

When comparing the macro-sequences of the six question topics, we face the problem that incomplete macro-sequences are very unevenly distributed over the topics, as appears in Table 3. The topics Gardening and Cycling have much more incomplete macro-sequences than, e.g., the topics Walking and Light Household Work.

The simple base question “Do you do X?” hardly ever requires repair activities by the interviewer, let alone partial questioning. So if a certain topic leads to many incomplete sequences, with only the base question asked and—negatively—answered, it would seem that this topic generally does not require much control activities of the interviewer. For a meaningful comparison of the macro-sequences belonging to different topics it is therefore necessary to restrict the comparisons to the complete macro-sequences only.

9 Three types of complete macro-sequences

Whether the additional turns in a relatively long complete macro-sequence are generated by a partial questioning strategy of the interviewer or not, can be seen from the coding of the second coding variable of the individual turns, mentioned in Box 2. For the topics light and heavy household work, the presence of code ‘h’ indicates that partial questioning has occurred. For the other four topics, partial questioning is indicated by a repetition of the code-string f–d, that is, a frequency question followed by a duration question (for one component of the question), and again succeeded by a frequency and a duration question for another component. If neither code ‘h’, nor the code string f–d. . .f–d, does occur in a relatively long macro-sequence, we may conclude that only repair activities have led to the additional turns.

By this reasoning we end up with three different types of complete macro-sequences:

Type A: relatively short macro-sequences (the AT-score is maximally 3) with only open loop control;

Type B: relatively long macro-sequences (AT more than 3) without partial questioning, so only with open loop control and repair activities; and

Type C: relatively long macro-sequences (AT more than 3) with partial questioning and open loop control, and possibly also repair activities.

Table 4 The distribution of the complete macro-sequences over the three types of control per question topic

Topic	Type A (%)	Type B (%)	Type C (%)	Total (%)	N
Walking	11	71	18	100	202
Cycling	28	57	15	100	107
Gardening	66	30	4	100	73
Sport	30	64	6	100	140
Light HW	32	20	48	100	228
Heavy HW	34	8	58	100	181
Total	30	40	30	100	931

$$\chi^2 = 334.9; \text{df} = 10; p < 0.001$$

10 The relationship between type of control and question topic

Of all 931 complete macro-sequences, 278 (30%) belong to Type A (only open loop control), 376 (40%) belong to Type B (open loop control plus repair activities) and 277 (30%) are of Type C (open loop control plus partial questioning and possibly repair activities). See Table 4.

There appear striking differences between the topics of the questionnaire with respect to the kind of control exercised by the interviewers.

Open loop control seemed to work quite well with the topic of ‘gardening’ (66%), resulting in generally short macro-sequences.

Feedback control, in the form of repair activities, is very much exercised with the topic of ‘walking’ (71%) mainly because the definition of ‘walking’ is not quite clear to the respondent (and sometimes not even to the interviewer). And the topics ‘sport’ and ‘cycling’ also require a lot of repair activities by the interviewer.

Feedforward control in the form of partial questioning is especially used with the topics ‘light and heavy household work’ (48%, respectively 58%). The questions about these topics are difficult to answer, and it is quite easy to decompose the questions about these topics into questions about separate components, i.e., separate household tasks.

11 The interviewer and the type of control

The exercise of partial questioning is certainly not scripted; it is based on the initiative of individual interviewers who think that decomposing the questions will lead to better results. We thus may expect that the degree to which partial questioning is used will differ between interviewers. The data of Table 5 give strong support to that expectation.

Table 5 Distribution of all complete macro-sequences over type of control, by interviewer number

Interviewer number	Type A (%)	Type B (%)	Type C (%)	Total (%)	N
#1	29	48	23	100	258
#2	23	27	50	100	248
#3	50	30	20	100	110
#4	30	49	30	100	315
All interviewers	30	40	30	100	931

$$\chi^2 = 90.8; \text{df} = 6; p < 0.001$$

Table 6 Type of control exercised by the interviewers in different phases of their fieldwork (begin, middle, and end); for complete macro-sequences only

Type of Control	Begin	Middle	End	Total
Type A: Open loop control only	25%	31%	32%	30%
Type B: Open loop control with repair	47%	39%	36%	40%
Type C: Using partial questioning	28%	29%	32%	30%
Total	100%	100%	100%	100%
Number of complete macro-sequences	273	370	288	931

$$\chi^2 = 7.3; df = 4; p = 0.12 \text{ (n.s.)}$$

Interviewer #2 used partial questioning (Type C) in 50% of 'her' macro-sequences. The three other interviewers did not differ very much with respect to the type of control used in their interviews. Interviewer #3 used in half of 'her macro-sequences' only the prescribed 'open loop control' in 'her' macro-sequences, therewith staying closest to the instructions of the researcher.

In order to see whether the control behavior of the interviewers changed during their fieldwork, the interviews were divided into three groups: the interviews held in the beginning of the fieldwork, those in the middle and those at the last part of the fieldwork.

From Table 6 it appears that during the course of the fieldwork there is a slight decrease of repair activities (Type B), compensated by an increase of both open loop control and partial questioning (Types A and C). However, these differences are not statistically significant. There might be some form of 'learning effect' during the fieldwork, but this effect is not very strong.

12 Summary and discussion

In the research project described in this paper, elderly respondents (65+) were asked detailed retrospective questions about 'mundane' physical activities like walking, cycling and household tasks. On the basis of methodological literature concerning the difficulty of retrospective questions (especially those with low saliency), and the more frequently occurring inadequate response behavior of elderly respondents (Schwarz et al. 1998), one would expect many 'no answers' and 'don't know answers'.

However, contrary to that expectation, the proportion of such inadequate answers appeared to be quite low, indicating that the interviewers must have done their very best to obtain adequate answers. For that purpose they, in principle, had three different methods at their disposal. Firstly, an optimal execution of the prescribed 'open loop control', that is, precisely following the text of the questionnaires and the instructions given by the researcher. Secondly, via the 'feedback loop control' of repair: if nevertheless inadequate answers are given, further probing is performed until eventually an adequate answer is obtained. And thirdly via 'feedforward loop control'; if the interviewer expects problems with posing this particular question to that particular respondent, she may adjust the question, for example by decomposing a general question into its components, making the question probably easier to answer (partial questioning).

To investigate to what degree feedback loop and feed forward loop controls are actually used by interviewers, the classical design of methods research of interviewer effects, i.e., the split ballot design, does not fit because this experimental design can only deal with the unidi-

Table 7 Summary of the relationships between the question topic, length and completeness of the corresponding macro-sequences, the causes of additional turns in the macro-sequences, and the type of control exercised by the interviewers

	Gardening	Sport	Cycling	Walking	Light HW	Heavy HW
High proportion of long macro-sequences				*	*	
High proportion of incomplete macro-sequences	*		*			
Causes of additional turns						
Definition problems		*		*		
Partial questioning/probing			*	*	*	*
Completion probing			*	*		
Confirmation and check						
Conferring of I and R						
Complex situation of R			*		*	*
Type of control by interviewer						
A: Open loop control only	*					
B: Feedback: repair		*	*	*		
C: Feedforward: partial questioning					*	*

rectional causal relation between interviewer behavior and resulting answers of respondents, but not with interviewer behavior that is the *result* of the expected or observed initial answers of the respondents. For our research purpose we therefore had to use another research design, i.e., a detailed analysis of the interactions between interviewer and respondent preceding the final response of the respondent.

In Table 7 a summary of the findings of our research is presented. From the analysis of question-answer sequences it appeared that the type of control exercised in these sequences, and as a consequence, the length of the macro-sequences, differs strongly by topic of the question. With ‘gardening’ the prescribed open loop control was sufficient, but the topic ‘walking’ required a lot of repair, whereas the topics of ‘light and heavy household tasks’ initiated much feedforward loop control in the form of partial questioning.

The differences regarding the behavior of interviewers and respondents between the six topics cannot be explained by differences in question format or question wording, because for all six topics the same question format and the same question wording were used. Neither can differences regarding interviewer characteristics explain the differences between the six topics as each interviewer posed questions about all topics. Nor can the differences between the six topics be explained by respondent characteristics as the respondents were about randomly assigned to the interviewers.

So we can conclude that question topics requiring more complex cognitive activities of the respondents, not only lead to more need for repair, but also to larger differences between the interviewers concerning their readiness to use feedforward loop control, in the form of partial questioning, in addition to, or as an alternative for, the open loop control as designed by the researcher.

Researchers who design questionnaires and train interviewers, would do well to reckon with the differences regarding the cognitive difficulties of topics of the questions. For quite difficult topics (like average time spent on irregular, but at the same time quite mundane activities of respondents), a different question format could be necessary, providing the respondent with clear information about the delineation of the topic, and with memory aids. Even a decomposition of the question topic into its components could be considered. In the

interviewer training ample time has to be devoted to learn interviewers to react adequately to inadequate initial responses of the respondents, i.e., to perform appropriate repair.

A detailed analysis of the interactions between interviewers and respondents during pilot interviews may provide the researcher with information about what topics need these extra efforts.

Acknowledgements This paper is based on data collected in the context of the Longitudinal Aging Study Amsterdam (LASA). This program is conducted at the Vrije Universiteit in Amsterdam and is supported by the Ministry of Health, Welfare, and Sports.

Appendix: Questions from the LASA physical activity questionnaire (LAPAQ) (Stel 2003)

Q1a. Do you **walk** outside? With walking outside we mean walking to go shopping or doing other daily activities, like visiting someone. We do not mean: a walking tour. 1. no (go to question Q2a); 2. yes.

Q1b. Did you walk during the past two weeks? 1. no (go to question Q2a); 2. yes.

Q1c. How many times did you walk during the past 2 weeks? ... times.

Q1d. How long did you usually walk each time? ... hours; ... minutes.

Q2a. Do you **cycle**? With cycling we mean cycling to go shopping or doing other daily activities, like visiting someone. With cycling we do not mean: a cycling tour. 1. no (go to question Q3a); 2. yes.

Q2b. Did you cycle during the last two weeks? 1. no (go to question Q3a); 2. yes.

Q2c. How many times did you cycle the past two weeks? ... times.

Q2d. How long did you usually cycle each time? ... hours; ... minutes.

Q3a. Do you have a **garden** (including allotment)? 1. no (go to question Q4a); 2. yes.

Q3b. Did you work in the garden during the past two weeks? 1. no (go to question Q4a); 2. yes.

Q3c. How many times did you work in the garden during the past two weeks? ... times.

Q3d. How long did you usually work in your garden each time? ... hours; ... minutes.

Q4a. Do you do **sports**? 1. no (go to question Q5a); 2. yes.

Q4c. How many times did you this sport during the past two weeks? ... times.

Q4d. How long did you usually do this sport each time? ... hours; ... minutes.

Q5a. Do you do **light household tasks**? With light household tasks we mean washing the dishes, dusting, making the bed, doing the laundry, hanging out the laundry, ironing, tidying up, and cooking meals. 1. no (go to question Q6a); 2. yes.

Q5c. How many days did you do light household tasks during the past two weeks? ... days.

Q5d. How long per day did you usually do light household tasks? ... hours; ... minutes.

Q6a. Do you do **heavy household tasks**? With heavy household tasks we mean window cleaning, changing the bed, beating the mat, vacuuming, washing or scrubbing the floor, and chores with sawing, carpeting, repairing or painting. 1. no (go to question Q7); 2. yes.

Q6c. How many days did you do heavy household tasks during the past two weeks? ... days.

Q6d. How long per day did you usually do heavy household tasks? ... hours; ... minutes.

References

- Cannel, C., Miller, P., Oksenberg, L.: Research on interviewing techniques. In: Leinhardt, S. (ed.) *Sociological Methodology* 1981, pp. 389–437. Jossey-Bass, San Francisco (1981)
- Deeg, D.J.H., Westendorp-de Serièrè, M. (eds.): *Autonomy and well-being in the aging population*. Report from the Longitudinal Aging Study Amsterdam 1992–1993. VU University Press, Amsterdam (1994)
- Dijkstra, W.: A new method for studying verbal interactions in survey-interviews. *J. Off. Stat.* **15**, 67–85 (1999)
- Dijkstra, W.: Transcribing, coding and analysing verbal interactions in survey interviews. In: Maynard, D. W., Houtkoop, H., Schaeffer N. C., van der Zouwen, J. (eds.) *Standardization and Tacit Knowledge; Interaction and Practice in the Survey Interview*, pp. 401–425. Wiley, New York (2002)
- Dijkstra, W.: Question–answer sequence. In: Best, S.J., Radcliff, B. (eds.) *Polling America; An Encyclopedia of Public Opinion*, pp. 612–616. Greenwood Press, Westport (2005)
- Dijkstra, W., Ongena, Y.: Question–answer sequences in survey-interviews. *Qual. Quant.* **40**, 983–1011 (2006)
- Graafmans, W.C., Bouter, L.M., Lips, P.: The influence of physical activity and fractures on ultrasound parameters in elderly people. *Osteoporosis Int.* **8**, 449–454 (1998)
- Groves, R.M.: *Survey Errors and Survey Costs*. Wiley, New York (1989)
- Groves, R.M., Fowler, F.J., Couper, M.P., Lepkowski, J.M., Singer, E., Tourangeau, R.: *Survey Methodology*. Wiley, Hoboken (2004)
- Houtkoop-Steenstra, H.: Conversational problems in telephone survey interviews. In: *Proceedings of the International Conference on Survey Measurement and Process Quality*, Bristol, April 1995, pp. 286–289. American Statistical Association, Alexandria VA (1995)
- Hyman, H.H., Cobb, W.J., Feldman, J.J., Hart, C.W., Stember, C.H.: *Interviewing in Social Research*. The University of Chicago Press, Chicago (1954/1975)
- Knaüper, B., Belli, R.F., Hill, D.H., Herzog, A.R.: Question difficulty and respondents' cognitive ability: the effect on data quality. *J. Off. Stat.* **13**, 181–199 (1997)
- Knaüper, B., Schwartz, N., Park, D.: Frequency reports across age groups. *J. Off. Stat.* **20**(1), 91–96 (2004)
- Moore, R.J., Maynard, D.W.: Achieving understanding in the standardized survey interview: repair sequences. In: Maynard, D.W., Houtkoop-Steenstra, H., Schaeffer, N.C., van der Zouwen, J. (eds.) *Standardization and Tacit Knowledge; Interaction and Practice in the Survey Interview*, pp. 281–312. Wiley, New York (2002)
- Ongena, Y.P., Dijkstra, W.: Methods of behavior coding of survey interviews. *J. Off. Stat.* **22**(3), 419–451 (2006)
- Rapoport, A.: Two approaches to general system theory. In: Geyer, R.F., van der Zouwen, J. (eds.) *Dependence and Inequality*, pp. 267–283. Pergamon Press, Oxford (1982)
- Schaeffer, N.C., Maynard, D.W.: From paradigm to prototype and back again: interactive aspects of cognitive processing in standardized survey interviews. In: Schwarz, N., Sudman, S. (eds.) *Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research*, Jossey-Bass, San Francisco (1996)
- Schwarz, N., Park, D., Knaüper, B., Sudman, S. (eds.): *Cognition, Aging, and Self-Reports*. Psychology Press, Philadelphia (1998)
- Stel, V.: *Prevention of fall accidents in older persons. From risk profile to intervention strategy*. Unpublished dissertation, Vrije Universiteit Amsterdam (2003)
- Stel, V., Smit, J.H., Pluijm, S.M.F., Visser, M., Deeg, D.J.H., Lips, P.: Comparison of the LASA physical activity questionnaire with a 7-day diary and pedometer. *J. Clin. Epidemiol.* **57**, 252–258 (2004)
- Tourangeau, R.: Cognitive science and survey methods. In: Jabine, T., Straf, M., Tanur, J., Tourangeau, R. (eds.) *Cognitive Aspects of Survey Design: Building a Bridge Between Disciplines*, pp. 73–100. National Academy Press, Washington (1984)
- Tourangeau, R., Rasinski, K.: Cognitive processes underlying context effects in attitude measurement. *Psychol. Bull.* **103**, 299–314 (1988)
- Tourangeau, R., Rips, L.J., Rasinski, K.: *The Psychology of Survey Response*. Cambridge University Press, Cambridge (2000)
- van der Zouwen, J., Smit, J.H.: Control processes in survey interviews: a cybernetic approach. *Kybernetes, Int. J. Syst. Cybern.* **34**(5/6), 602–616 (2005)
- van der Zouwen, J., Smit, J.H.: Effective repair in survey interviews. *Qual. Quant.* **40**, 245–258 (2006)
- van der Zouwen, J., Smit, J.H., van der Horst, M.H.L.: Reporting the frequency and duration of household tasks by elderly respondents: the effect of different interview strategies on data quality. *Proceedings of the 60th AAPOR Annual Conference*, May 12–15, 2005, Miami Beach, FL, pp. 4007–4014. American Statistical Association, Alexandria [CD-ROM] (2005)
- Wiener, N.: *Cybernetics*, 2nd edn. Wiley, New York (1948/1965)