Suggestive Interviewer Behaviour in Surveys: An Experimental Study[1]

Johannes H. Smit[2]*, Wil Dijkstra*[3]*, and Johannes van der Zouwen*[3]

The consequences of suggestive interviewer behaviour as a potential source of bias in obtaining valid answers in survey settings are discussed. It is hypothesized that: (1) suggestive interviewer behaviour while asking closed questions, or during probing, influences the responses and their distributions; and (2) parameter estimations of relationships with variables measured with questions influenced by suggestive behaviour are affected too. Three kinds of more or less suggestive interviewer behaviour concerning the presentation of response alternatives, following a closed question about consequences of aging, were systematically varied in a field experiment across different groups of randomly selected older (55+) respondents ($N = 235$). After obtaining a response to the question, the interviewer asked for any reasons for that particular response, thereby systematically suggesting a particular aspect of aging. After these manipulations, respondents were asked to evaluate a number of aspects, among them those previously suggested to the respondent. The distributions of the responses to the closed question proved to differ between experimental groups: suggested answers were indeed mentioned more often ($p < .001$). Suggestive probing had an effect too: one of the suggested aspects was evaluated as having greater effect than aspects that were not suggested ($p = .035$). Finally, the correlation between the responses to the closed question and another variable, age, turned out to be dependent on the experimental condition, with correlation coefficients ranging from $r = .03$ to $r = .35$. The experiment shows that suggestive interviewing indeed affects the quality of the data collected.

Key words: Suggestive questions; suggestive probing; field experiment.

1. Introduction

In survey research, standardized interviews are frequently used, aiming at valid responses to the questions of the questionnaire. However, inadequate interviewer behaviour may cause the respondent to answer the questions in a way he or she would not have done had the interviewer not behaved in a particular manner. This difference belongs to the general category of ''interviewer effects.'' Interviewer effects may influence both the answers and their distributions, as well as the relations between this variable and other variables.

In studies on interviewer effects, two approaches are usually followed (Groves 1989). The first approach studies the relation between particular interviewer *characteristics*

[2] Department of Sociology and Social Gerontology, Vrije Universiteit, De Boelelaan 1081C, 1081 HV, Amsterdam, The Netherlands.
[3] Department of Social Research Methodology, Vrije Universiteit, De Boelelaan 1081C, 1081 HV, Amsterdam, The Netherlands.

(like age, sex, and education level) and the respondents' answers (e.g., Groves and Fultz 1985). The second approach studies the relation between (inadequate) interviewer *behaviour* (e.g., incorrect reading of questions, not following instructions, suggestive questioning or suggestive probing) and the answers obtained (e.g., Van der Zouwen, Dijkstra, and Smit 1991). In this article we focus on the second approach, and discuss the consequences of two types of inadequate interviewer behaviour for the quality of the data, i.e., suggestive questioning and suggestive probing.

1.1. Previous research on suggestive questioning

In methodological literature it is generally stated that questions have to be posed in a neutral way (e.g., Foddy 1993). That is, without giving cues about which response is expected from the respondent, and without making one of the response alternatives more salient to the respondent than the other ones.

However, in empirical studies of the effects of suggestive questioning, rather divergent findings are obtained. Richardson, Dohrenwend, and Klein (1965, p.186) conducted an experiment and reported that ''the results show that responses to leading questions contain no more distortions than responses to nonleading questions.'' Dohrenwend (1970, p.174) found in an experimental study that ''directive questions can be used to increase the reporting of controversial behaviour.'' On the other hand, Loftus and Zanni (1975) found that directive questioning had a negative effect on the quality of respondents' reports.

These contradictory results stimulated us to design and conduct an experiment specifically aimed at the assessment of the effects of suggestive questioning and suggestive probing.

1.2. Suggestive questioning

A distinction can be made between questions, deliberately worded suggestively by the researcher, and suggestive formulations of the questions as posed by the interviewers during the question–answer interaction. In the first case, the researcher assumes that the suggestive formulation of the question will stimulate the respondent to give valid answers, for example, because this formulation is assumed to neutralize the effects of any ''social desirability bias'' (Kinsey, Pomeroy, and Martin 1948; Dohrenwend 1970). In the second case, the interviewer formulates questions suggestively, obviously without approval of the researcher.

In an explorative study of verbatim protocols of 654 question–answer sequences, we found that in more than one-quarter of these sequences, alternatives were presented in a suggestive manner. This behaviour especially occurred when somewhere in the interaction something went wrong, and interviewers tried to repair this ''mistake.'' Flaws in the inter-action mainly originated when the respondent formulated an answer before the (complete) set of alternatives was presented.

There may be two reasons for this type of respondent behaviour. Firstly, the respondents are not aware that they should answer by mentioning only one from a limited set of response alternatives, and they answer the question before the interviewer has had a chance to present the complete set of alternatives. Secondly, the interviewer deliberately waits for the respondents' answers after having presented only a selection of the

alternatives from the complete set. The interviewer probably makes the selection based on the respondent's previous answers.

1.3. Suggestive probing

Suggestive probing (i.e., suggestively asking for the reasons for the choice of a particular response alternative) is also often mentioned in the methodological literature as a form of commonly occurring inadequate interviewer behaviour (Cannell, Oksenberg, and Converse 1979; Brenner 1982; Sykes and Collins 1992).

   Explorative analyses of 317 transcribed question–answer sequences in which probing behaviour occurred showed that in nearly 25% of these question–answer sequences probes were formulated suggestively. Typically, these suggestive probes were formulated as ''yes/no questions'' in which a specific ''candidate response'' was offered to the respondent.

1.4. An example

In the next question–answer sequence both types of inadequate interviewer behaviour are illustrated. The questionnaire contains the following questions:

1. *How satisfied are you with your income?*
   (1. Very satisfied; 2. Satisfied; 3. Neither satisfied nor dissatisfied; 4. Dissatisfied; 5. Very dissatisfied).
   Interviewers were instructed to read aloud all five alternatives.
2. *Why are you satisfied/dissatisfied?*
   Instruction to interviewer:
   Probe further for the reasons for the answer.
   Make detailed notes of the answer.

In one of the interviews the following type of interaction evolved:

1. *Interviewer:* How satisfied are you with your income?
2. *Respondent:* Well, it's not easy but I am not complaining either.
3. *Interviewer:* What if you had to choose between ''dissatisfied'' and ''very dissatisfied?''
4. *Respondent:* In that case ''dissatisfied.''
5. *Interviewer:* Are you dissatisfied because you are having a hard time making ends meet?
6. *Respondent:* Well, my pension does not make me rich. Most of the time I am trying to make ends meet. I cannot afford to do nice things. But I guess my situation is not exceptional.

Comments:
1. The interviewer reads the question as worded in the questionnaire.
2. The respondent gives an answer before the interviewer has had the opportunity to present the five response alternatives.
3. Based on that answer, the interviewer presents a reduced set of response alternatives, thereby ''repairing'' the first part of the interaction.

4. The respondent selects one alternative from the reduced set.
5. Instead of using the wording of the probe question from the questionnaire, the interviewer formulates a probe by him/herself. This probe leads the answer of the respondent in a certain direction (suggestive probe).
6. The respondent formulates an answer in the direction suggested by the interviewer.

2. Methodology

2.1. Sample

To reveal the influence of suggestive interviewer behaviour, the presence of suggestive questioning should be varied systematically over different groups of respondents. This allows comparison of the answers (distributions) of respondents to whom a particular answer was suggested with answers of respondents to whom no suggestions were made. To investigate the effects of suggestive interviewer behaviour, we conducted a field experiment to study both the effects of suggestive presentation of answer alternatives and the effects of suggestive probing.

A random sample ($N = 345$), stratified by age and sex, was drawn from the municipal register of Sassenheim, a small town located in the western part of The Netherlands. All sample persons were 55 to 89 years of age. The interview was introduced as an investigation of the independence and health of elderly people. The sample persons received a letter from the Vrije Universiteit informing them about the purpose of the study. They were promised a small reward if they participated in the survey. Afterwards an interviewer contacted them to make an appointment for an interview. All interviews were held in the respondents' homes, using CAPI (Computer Assisted Personal Interview) procedures. The field work was conducted from January 1992 to April 1992.

2.2. Manipulation 1: suggesting answer alternatives

The first experimental manipulation concerned the suggestion of answer alternatives. The aim was to determine the effect of ''repairing'' the interaction by the interviewer, as illustrated in Section 1.4. To simulate such a flaw in the interaction, the interviewers were instructed to read the first part of the question and then wait for a first reaction (i.e., answer) of the respondent, after which they had to present one or more of the response alternatives.

The question read:

''Did you, over the last years, have to adapt to the fact that you are getting older?''
The answer alternatives are:
1. Very often; 2. Often; 3. Now and then; 4. Not often; and 5. Never.

It should be noted that the question itself was formulated as a yes/no question. A yes/no question in combination with a set of answer alternatives cannot be viewed as good survey practice – but this practice is not uncommon. Our procedure is used to enhance the chance of ''inadequate'' responses which the interviewer then attempts to repair. The induced flaw in the interaction was repaired in three different ways. In the first experimental condition, interviewers were instructed to present a set of answer alternatives which

included the first answer of the respondent. For example: if the answer tends to be positive (''Yes,'' ''I guess so,'' etc.), the interviewer should ask: ''Has this occurred ''very often,'' ''often,'' or ''now and then?'' If the answer tends to be negative (''No,'' ''Actually not,'' etc.), the interviewer should present: ''Has this occurred ''now and then,'' ''not often,'' or ''never?'' If the answer tends to be ''in between'' (''Sometimes,'' ''At times,'' etc.), the interviewer was instructed to ask: ''Has this occurred ''often,'' ''now and then'' or ''not often?''

In the second condition the interaction flaw was repaired by only presenting either the alternative ''often'' or the alternative ''not often,'' depending on the direction of the first answer of the respondent. Finally, in the third condition, the interaction flaw was repaired by the presentation of only the answer alternative, the interviewer considered to fit best with the previously given answer.

In all three conditions the original set of response alternatives is restricted to a smaller one. In the first condition three out of five alternatives are presented; in the second and third condition only one alternative is presented to the respondent: ''Is it X?'' In both last conditions the *direction* of the previously given response is taken into account while selecting X. But in the second condition, in contrast to the third one, this selection is not based on information about the *intensity* given in the previous response. Condition one is closest to a neutral presentation of the question, condition two is considered to be the most suggestive one: It restricts the response repertory of the respondent and it does not fully utilize relevant information already provided by the respondent. For the sake of brevity we will indicate condition 1 by using the term ''neutral,'' condition 2 by the term ''suggestive,'' and condition 3 by ''restricted.''

2.3.   *Manipulation 2: suggestive probing*

Methodological research has shown that responses to survey questions are affected by the way preceding questions are formulated and answered (Schwarz and Sudman 1992). Previous questions may induce, selectively, the saliency of certain topics, thereby enhancing the probability that these topics will be mentioned in subsequent questions.

In our field experiment, we wanted to investigate the possible influence of suggestive probing related to the question about the adaptation to aging (mentioned in the previous section) on the response to a following question. Three experimental conditions were distinguished. In the first experimental condition, interviewers were instructed to use only neutral probes like ''Can you tell me something more about this?'' or ''Will you please clarify your answer?'' In the second condition, interviewers were instructed to probe for adaptations to aging, related to the financial situation of the respondent. In the third condition, the interviewers were instructed to probe for adaptations related to contacts with family and friends.

After responding to the probing questions, all respondents were asked to indicate whether getting older affected a number of aspects of their lives, including their financial position and contacts with family and friends. This question was worded as follows:

''How would you rate the influence of ''growing older'' on each of the following aspects?'' Choose a number between 0 and 100; 0 means: no influence at all; 100 means: a very strong influence.

The aspects presented were:

1. Your day to day duties; 2. Your financial situation; 3. Your contacts with family or friends; 4. Your holidays; 5. Your involvement in current affairs.

It was expected that respondents exposed to a probe question related to a particular aspect (i.e., Your financial situation, in condition 2, and Your contacts with family or friends in condition 3) would subsequently evaluate this aspect as being more salient than respondents probed in a more neutral way (condition 1). The difference in saliency would then result in a higher score for the suggestively probed aspect.

2.4. *Selection and training of the interviewers*

Interviewers were recruited via newspaper advertisement and bulletin boards at the University of Leiden. Important selection criteria included: higher education, affinity with elderly people, and social skills. Ten out of 39 applicants were selected. Most of them were psychology graduates, their ages ranging from 22 to 39, and all female. The interviewers were informed that the study included an experiment.

The training of the interviewers took place during five sessions, each lasting six hours. Video examples illustrated basic interview rules; role playing was used to practise interviewer skills and handling the experimental manipulations. Each interviewer conducted a test interview which was discussed afterwards.

Each interview was audio-taped. These tapes made it possible to monitor interviewer behaviour. During the fieldwork, interviewers were closely supervised. Two additional training sessions took place in which the experimental manipulations were practised again. Mistakes and inadequacies were discussed with the interviewers. The interviewers received advice on improving their performances. By following this procedure it was expected that interviewer performance regarding the experimental conditions would improve.

The ten interviewers were randomly assigned to the three experimental conditions: four interviewers were assigned to condition 1 (neutral for suggesting answer alternatives, and neutral for suggestive probing), three interviewers to condition 2 (suggestive for suggesting answer alternatives, and financial for suggestive probing), and three interviewers to condition 3 (restricted for suggesting answer alternatives, and social contacts for suggestive probing). Apart from the experimental conditions, the questionnaire was the same for every interviewer. The respondents were randomly assigned to the interviewers.

2.5. *Hypotheses*

We formulated three hypotheses about the effects of suggestive questioning and suggestive probing.

Hypothesis 1:
The answers to the question ''Did you, over the last years, have to adapt to the fact that you are getting older?'' are different for the three experimental conditions, with the largest differences between conditions 1 and 2.

Suggestive questioning may not only affect the answers (distributions) per se, but also

their relationship to other variables. It is quite conceivable that ''adaptation to getting older'' is related to age itself. Hence we state:

Hypothesis 2:
The strength of the relationship between the answer to the question ''Did you, over the last years, have to adapt to the fact that you are getting older?'' on the one hand, and the age of the respondent on the other, differs per condition, with the largest differences between conditions 1 and 2.

Hypothesis 3 reads:
Respondents probed for adaptations related to their financial situation, and their contacts with family and friends will assign a higher score to the item regarding ''financial aspects'' and ''social contacts'' than respondents probed in a neutral way.

3. Results

3.1. Response/nonresponse and manipulation checks

In total 235 interviews were completed (68% of the original sample of $N = 345$). Most nonresponse was caused by refusal (19%) or ineligibility (e.g., illness 9%). No significant differences between the age- and sex-distributions of respondents in each of the three experimental conditions were found.

Item nonresponse did not differ across the experimental conditions. Respondents with item nonresponse on an experimental manipulation were not included in the analysis.

Verbatim protocols of 100 interactions were evaluated for their adequacy by judges who were versed in the different manipulations. In ten out of the 100 reviewed interactions, the manipulations appeared to have failed.

3.2. Effects of the suggestive presentation of response alternatives

To test hypothesis 1, answer distributions for the three conditions in this manipulation

Table 1.   Response distributions of respondents assigned to one of the three experimental conditions related to the adaptation question (per cent)

Experimental condition	Very often	Often	Now and then	Not often	Never	Total
Condition 1	12	19	46	6	17	100 ($n = 89$)
Condition 2	9	33	19	25	15	100 ($n = 64$)
Condition 3	4	20	37	17	23	100 ($n = 71$)

Condition 1: neutral, the three alternatives presented encompass the previous response.
Condition 2: suggestive, only the alternatives often or not often are presented, dependent on the previous answer.
Condition 3: restricted, only the alternative which is closest to the previous answer is presented.
Due to item nonresponse the total number of respondents is less than 235.

Results of the likelihood-ratio test:

Condition (1, 2, 3)	L-test $= 27.27$	$df = 8$	$p < .001$
Condition (1, 2)	L-test $= 22.15$	$df = 4$	$p < .001$
Condition (1, 3)	L-test $= 9.90$	$df = 4$	$p = .044$
Condition (2, 3)	L-test $= 9.51$	$df = 4$	$p = .049$

were compared using a likelihood-ratio test (*L*-test). The *L*-test verified whether there are differences between the distributions of the answers to the ''adaptation'' question.

From Table 1 it appears that there are substantial differences between answer distributions of respondents belonging to each of the three conditions. In condition one, next to two other alternatives, interviewers always presented the alternative ''now and then.'' In this condition the alternative now and then was mentioned most frequently by the respondents. In condition two, interviewers only presented the alternatives ''often'' or ''not often.'' These two alternatives were most frequently mentioned by respondents interviewed in this condition. Apparently, the interviewers' suggestions were often adopted by respondents. In condition three, interviewers themselves had to decide which alternative to present to the respondent. Noticeable is that in condition three the alternative ''very often'' was least mentioned by the respondents. It might be that considerations of social desirability prevent interviewers from suggesting this alternative. In agreement with hypothesis 1, the differences between the three conditions are largest for conditions 1 and 2, the neutral versus the suggestive presentation of the response alternatives.

Next, the effect of suggestive interviewer behaviour on the relationship with another variable was examined, i.e., the relation between the response to the question on ''adaptations'' and the age of the respondent. By means of regression analysis the relation between these two variables was studied for each of the three conditions. For all respondents taken together, a positive correlation ($r = .15$) was established. Within condition two this correlation appeared to be much stronger than in the other conditions (see Table 2). In agreement with hypothesis 2 the differences with respect to the strength of the relationship between these two variables are largest for conditions 1 and 2.

We may draw the conclusion that the data from our field experiment support hypotheses 1 and 2. Suggestive presentation of answer alternatives not only influences the responses but also the relationship with another variable, i.e., the age of the respondent. This outcome is obviously consequential for the testing of relations with other variables.

3.3.   *Effects of suggestive probing*

According to hypothesis 3, suggestive probing will affect the scores regarding the saliency of the items ''financial aspects,'' and ''contacts with family and friends,'' items which were presented after the probing manipulation.

To test this hypothesis, scores on the item ''financial aspects,'' and ''social contacts''

Table 2.   *The correlation between the answer to the adaptation-question and the age of the respondent, for each experimental condition*

Exp. condition	*r*	*p*-value	*N*
Condition 1	.03	.741	88
Condition 2	.35	.005	63
Condition 3	.13	.284	70
Total	.15	.029	221

Condition 1: neutral, the three alternatives presented encompass the previous response.
Condition 2: suggestive, only the alternatives often or not often are presented, dependent on previous answer.
Condition 3: restricted, only the alternative which is closest to the previous answer is presented.
Due to item nonresponse the number of respondents is less than 235.

given by respondents assigned to condition one (neutral probing), were compared with scores of respondents assigned to the other conditions (i.e., deliberately probing for financial aspects, and social contacts, respectively). It was expected that respondents belonging to condition two (or three) would score higher on the item financial aspect (or social contacts) than respondents probed in a more neutral way. Because the answers were not normally distributed, a Mann-Whitney-Wilcoxon test, instead of the usual *t*-test, was used to test the difference between the response distributions. For the item financial aspects both median scores were equal ($Me = 10$). For the item social contacts a significant difference ($p = .035$) in the expected direction was obtained: in the neutral condition the median score for this item was 15, whereas in the condition where social contacts were suggested, a median score of 50 was obtained. So the data partly support hypothesis 3.

4. Summary

The purpose of the field experiment was to determine the effects of suggestive interviewer behaviour on the quality of the answers provided by respondents in survey interviews. Suggestive questioning not only affects the response distribution of the variable concerned, but also the strength of the relationship between different variables. In one of the conditions included in our field experiment, an effect of probing behaviour was found: suggestive probing for the aspect contacts with family and friends significantly influenced the answers to the subsequent question concerning the effect of aging on these contacts. Thus suggestive interviewing indeed affects the quality of the data collected.

5. References

Brenner, M. (1982). Response-Effects of ''Role-restricted'' Characteristics of the Interviewer. In Response Behaviour in the Survey Interview, W. Dijkstra and J. van der Zouwen (eds). London: Academic Press.

Cannell, C.F., Oksenberg, L., and Converse, J.M. (1979). Experiments in Interviewing Techniques: Field Experiments in Health Reporting, 1971–1977. Ann Arbor, Michigan: Institute for Social Research, The University of Michigan.

Dohrenwend, B.S. (1970). An Experimental Study of Directive Interviewing. Public Opinion Quarterly, 34, 117–125.

Foddy, W. (1993). Constructing Questions for Interviews and Questionnaires. Cambridge: Cambridge University Press.

Groves, R.M. (1989). Survey Errors and Survey Costs. New York: John Wiley.

Groves, R.M. and Fultz, N.H. (1985). Gender Effects Among Telephone Interviewers in a Survey of Economic Attitudes. Sociological Methods & Research, 14, 31–52.

Kinsey, A.C., Pomeroy, W.B., and Martin, C.E. (1948). Sexual Behavior in the Human Male. Philadelphia: Saunders.

Loftus, E.F. and Zanni, G. (1975). Eyewitness Testimony: The Influence of the Wording of a Question. Bulletin of the Psychonomic Society, 5, 86–88.

Richardson, S.A., Dohrenwend, B.S., and Klein, D. (1965). Interviewing, Its Forms and Functions. New York: Basic Books.

Schwarz, N. and Sudman, S. (eds) (1992). Context Effects in Social and Psychological Research. New York: Springer-Verlag.

Sykes, W. and Collins, M. (1992). Anatomy of the Survey Interview. Journal of Official
    Statistics, 8, 191–207.

Van der Zouwen, J., Dijkstra, W., and Smit, J.H. (1991). Studying Respondent-Interviewer
    Interaction: The Relationship Between Interviewing Style, Interviewer Behavior, and
    Response Behavior. In Measurement Errors in Surveys, P.P. Biemer, R.M. Groves,
    L.E. Lyberg, N.A. Mathiowetz, and S. Sudman (eds). New York: John Wiley.