

A comparison of response effects in self-administered and telephone surveys

Bishop, George E.; Hippler, Hans-Jürgen; Schwarz, Norbert; Strack, Fritz

Veröffentlichungsversion / Published Version

Forschungsbericht / research report

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Bishop, G. E., Hippler, H.-J., Schwarz, N., & Strack, F. (1987). *A comparison of response effects in self-administered and telephone surveys*. (ZUMA-Arbeitsbericht, 1987/11). Mannheim: Zentrum für Umfragen, Methoden und Analysen - ZUMA-. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-66484>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

A Comparison of Response Effects in
Self-Administration and Telephone
Surveys

George F. Bishop
Hans-Jürgen Hippler, Norbert Schwarz
Fritz Strack

ZUMA-Arbeitsbericht Nr. 87/11

Zentrum für Umfragen, Methoden und Analysen
e.V. (ZUMA)
Postfach 12 21 55

D-6800 Mannheim 1

Seit Juli 1983 sind die ZUMA-Arbeitsberichte in zwei Reihen aufgeteilt:

Die **ZUMA-Arbeitsberichte** (neue Folge) haben eine hausinterne Begutachtung durchlaufen und werden vom Geschäftsführenden Direktor zusammen mit den übrigen Wissenschaftlichen Leitern herausgegeben. Die Berichte dieser Reihe sind zur allgemeinen Weitergabe nach außen bestimmt.

Die **ZUMA-Technischen Berichte** dienen zur hausinternen Kommunikation bzw. zur Unterrichtung externer Kooperationspartner. Sie sind nicht zur allgemeinen Weitergabe bestimmt.

The attached reprint replaces ZUMA-Arbeitsbericht No. 87/11 by the same authors.

Bishop, George, Hippler, Hans-J., Schwarz, Norbert, & Strack, Fritz. A comparison of response effects in self-administered and telephone surveys. In R.M. Groves, P. Biemer, L. Lyberg, J.T. Massey, W.L. Nicholls, & J. Waksberg (Eds.), Telephone survey methodology, New York: Wiley, 1988.

Table 7. Age Mode Effects on a Question Battery

	Very Likely	Fairly Likely	Fairly Unlikely	Very Unlikely
Age 65+				
Face to face	11%	32%	27%	29%
Telephone, read out	22	30	17	30
Absolute difference	11	2	10	1
Aged under 65				
Face to face	14	35	29	22
Telephone, read out	18	36	25	21
Absolute difference	4	1	4	1

CHAPTER 20

A COMPARISON OF RESPONSE EFFECTS IN SELF-ADMINISTERED AND TELEPHONE SURVEYS

George F. Bishop

Institute for Policy Research, University of Cincinnati

Hans-Juergen Hippler and Norbert Schwarz

Center for Surveys, Methods, and Analysis

Fritz Strack

*University of Mannheim*¹

Many experiments have shown that the results of social surveys can be significantly affected by the way in which the questions are worded, the form in which they are presented, and the order or context in which they are asked. Nearly all of this evidence, however, has come from survey interviews conducted either face to face or by telephone (see Bishop 1982, 1983, 1984, 1985, 1986, 1987; Kalton et al., 1978, 1980; Krosnick and Alwin, 1987; Schuman et al., 1983, 1986; Schuman and Ludwig, 1983; Schuman and Presser, 1981; Smith, 1987; see also Dijkstra and van der Zouwen, 1982). With one exception (Hippler and Schwarz, 1986), none of the well known response effects in this literature have, to our knowledge, been replicated in a self-administered or mail questionnaire, in which respondents typically have more time to think about each question and the implications of their answer to one question for their answer to

¹The authors would like to thank Paul Biemer, Edith Desiree de Leeuw, and Johannes van der Zouwen for their comments and suggestions on revising the previous draft of this chapter.

another.² It would thus be useful to know which response effects generalize to self-administered surveys, which do not, and why.³

Theoretically, we should expect some response effects to occur in both self-administered surveys and telephone or face to face interviews, but others should either disappear or become much less pronounced in magnitude in a self-administered situation. Question order effects, for example, should either vanish or become negligible in a self-administered survey because, unlike in a telephone or face to face interview (without a show card), respondents can consider all the questions and response alternatives before answering. So, for the same reason, should most response order effects be reduced in magnitude, if not eliminated, in a self-administered instrument, with the exception of long lists or item scales that respondents may inspect too hastily (see Schuman and Presser, 1981, pp. 72-74). In contrast, question form and wording effects should be just as likely to occur in a self-administered survey as in a telephone or face to face interview because the information presented to respondents in all of these modes of data collection (e.g., the presence or absence of a middle response alternative; the word "forbid" or "allow") is essentially equivalent. Because respondents in a self-administered survey have more time to think about the meaning of the questions, however, subtle variations in how they interpret them may well occur, resulting in significant differences between this and other modes of data collection.

To test these hypotheses, we designed a cross cultural experiment and replication that compared the response effects of variations in question form, wording, and context in a telephone survey with those in a self-administered survey, the two modes of data collection that we thought

²Since drafting this chapter, we have learned of a recent study by Ayidiya (1987), suggesting that one well known question order effect involving the Communist/American newspaper reporter items used by Schuman and Presser (1981) does replicate in a mail survey, but that another involving the same abortion item used in our experiment does not, confirming the findings shown in Table 3 below. The data from Ayidiya's study also show that a variety of recency response order effects reported in the literature are either eliminated or significantly reduced in magnitude in a mail survey, confirming too the results presented in Table 1 below. Furthermore, Ayidiya has found that acquiescence response effects due to question format are just about as likely to occur in a mail survey as in a telephone or face to face interview, confirming our findings on question form effects as well. So there is now convergent evidence for our hypotheses from an independent investigation.

³There are, of course, other inherent differences between self-administered and telephone or face to face surveys (e.g., interviewer effects, task related effects, etc.) which might account for some of the variation in response effects by mode of data collection, but we suspect that they are relatively minor. Though we have no conclusive evidence that mail survey respondents tend to look ahead at the questions and answers, this assumption would seem quite plausible.

were most different from one another.⁴ The experiment and the replication were done at about the same time, with the same questions, with similar populations: college students at the University of Cincinnati in the United States and students at the University of Mannheim in West Germany. The principal reason for using student subjects, other than limited resources, was that we did not think a sufficiently high response rate could have been achieved with a self-administered survey of the general public, whether done by mail or face to face delivery. Any differences we might find between the two modes of data collection with the general public might then be due to a difference in response rates. The use of West Germany and the United States as cultural settings for the experiment was purely a matter of convenience, an outgrowth of a visit by the first author to the survey research center, ZUMA, in Mannheim. Though the authors assumed cultural differences between the two societies would affect the marginal distribution of responses to various questions (e.g., a more "conservative" pattern in the United States), they expected that response and mode of data collection effects, if they were truly universal, should replicate from one cultural setting to another.

1. RESEARCH DESIGN

1.1 The Experiment in the United States

The data for this experiment were collected in February, March, and April of 1986 from a systematic random sample of 724 graduate and undergraduate students selected from a current telephone directory for the University of Cincinnati.⁵ Half of these students were randomly assigned to be interviewed by telephone; the other half received the questionnaire in a self-administered form that was personally delivered to their residence and returned either by mail or by having it picked up by

⁴Our assumption is that respondents have the least amount of time to think about the question and their answer in a telephone interview, somewhat more time to stop and think in a face to face interview, and the most time in a self-administered survey where they can look over all the questions and responses before answering. Thus we would expect to find the largest difference in response effects between telephone and self-administered surveys.

⁵We are unable to evaluate any bias due to nonresponse because the directory service on campus does not provide any population statistics or documentation on the accuracy of the directory. The response rate for the two surveys as reported below, however, was relatively high, indicating that nonresponse bias was probably not a significant source of error.

the person who delivered it, typically the latter.⁶ The response rate for the telephone survey was 83.9 percent; for the self-administered survey, with an intensive followup, it was 76.8 percent.⁷ In both the telephone and the self-administered surveys, respondents were randomly assigned to receive either form A or form B of the questionnaire (see Appendix).

Table 1 outlines the various experiments reported in this chapter.⁸ A majority of the items were exact replications of questions from previous split-ballot experiments conducted by Schuman and his associates or by Bishop and his coworkers. The authors constructed the question on left/right political identification as a substitute for an item on liberal/conservative identification used by Schuman and Presser (1981) because the latter terms did not have much currency in the political culture of West Germany, whereas the concepts of "the left" and "the right" seemed more comparable in meaning across the two cultures. Similarly, we designed the question about forbidding or allowing smoking in public places such as restaurants because we thought this issue would be somewhat more salient and comparable in meaning across the two societies than those used in previous investigations of the forbid-allow effect (Hippler and Schwarz, 1986; Schuman and Presser, 1981, pp. 276-283). Finally, we created the "International Trade Act of 1986" as a comparable, cross cultural surrogate for the fictitious "1975 Public Affairs Act" invented by Bishop and his colleagues (1980, 1986) and the real, but obscure, legislative acts used by Schuman and Presser (1981). So, as with the questions about left/right political identity and forbidding or allowing smoking in public places, the experiment with the International Trade Act

⁶The first author would like to thank Susan Ackerman, Shirley Frayer, and Andy Smith for their conscientious efforts in helping him to collect these data. He owes them all a big debt of gratitude.

⁷Most of the nonrespondents in both the telephone and the self-administered surveys were students who could not be located because the directory numbers listed for them had been disconnected. Only 10 potential respondents (3 men, 7 women) refused to be interviewed, all of them in the self-administered condition in which the questionnaire was to be personally delivered. Fourteen potential respondents in the self-administered condition did not return their questionnaires even after repeated reminders to do so by telephone. Otherwise, cooperation was unusually high compared to surveys of the general public.

⁸In addition, the questionnaire in both the German and the U.S. studies included two experiments on the effects of question length on responses to questions about respondents' memories of the wars in Afghanistan and the Falkland Islands. These data are part of a separate study on memory effects and so they are not included here. The questionnaires in both the self-administered and the telephone surveys in Cincinnati also included a question order experiment with items from the NORC General Social Survey concerning the death penalty and the harshness with which criminals are treated by the courts, as well as various measures of involvement with the death penalty issue. These items were part of an independent study by Susan Ackerman and are therefore not analyzed here.

item represents a conceptual, rather than an exact, replication of previous split-ballot studies.

1.2 The Replication in West Germany

The conditions under which the data were collected in the German setting were somewhat different, though the wording and sequence of the questions in this study were, with one minor exception, identical to those in the U.S. experiment.⁹ First of all, the data for the German study were collected during the first two weeks of April 1986, whereas the U.S. data were gathered in February and March as well as early April of 1986. To our knowledge, however, there were no international or domestic events in February and March of that year, nor for that matter in April, that might have differentially affected responses to any of the questions in either of the two experiments.

While the U.S. study included graduate and undergraduate students from a variety of disciplines, the subjects for the German investigation were all undergraduates at the University of Mannheim, majoring in law and business administration. Also, unlike the U.S. study, these students were initially contacted in the classroom and then asked to participate in a survey. After they agreed to participate, they were randomly assigned to one of the four conditions: the telephone or the self-administered survey and either Form A or Form B of the questionnaire. Subjects assigned to the self-administered condition received one of the two versions of the questionnaire and were asked to fill it out immediately, whereas subjects assigned to the telephone condition completed a one page questionnaire for a separate study in which they were asked on the second page for their telephone number, the best time for contacting them, and their first name. Most of the questionnaires were completed within 12 minutes. All of the questionnaires were then collected and each participant got a mechanical pencil as a reward.

During the following three days the subjects in the telephone condition were interviewed by five professional interviewers. Because of a concern over the relatively small sample sizes in the German experiment, the interviewers were instructed and trained to go through the interview

⁹ On the question about nuclear power plants (see Appendix), the U.S. version reads, "Some people say that the United States needs to develop *new* power sources from nuclear energy...[emphasis added]," whereas the English translation of the German version reads, "...*alternative* power sources [emphasis added]." We doubt that this minor change in wording made an important difference in the meaning of the item, though we really cannot be sure without performing still another experiment.

Table 1. List of Experiments in Question Form, Wording, and Context

	Form A	Form B
<i>Response Order</i> Divorce	Middle alternative presented in <i>last</i> position	Middle alternative presented in <i>middle</i> position
Nuclear power plants	Middle alternative presented in <i>last</i> position	Middle alternative presented in <i>middle</i> position
<i>Question Order</i> Abortion	Question on abortion if birth defect, asked <i>before</i> question about abortion if woman wants no more children	Question on abortion if birth defect, asked <i>after</i> question about abortion if woman wants no more children
Trade with Japan	Question about trade restrictions by Japan asked <i>before</i> question about trade restrictions by U.S./Germany	Question about trade restrictions by Japan asked <i>after</i> question about trade restrictions by U.S./Germany
<i>Middle Response Alternative</i> Marijuana penalties Left/right pol. ID Defense spending	Middle alternative omitted Middle alternative offered <i>Middle</i> alternative offered	Middle alternative offered Middle alternative omitted <i>No opinion alternative</i> offered instead of middle alternative
<i>No Opinion Alternative</i> Arab/Israel relations International Trade Act of 1986	No opinion alternative offered No opinion alternative omitted	No opinion alternative omitted No opinion alternative offered
<i>Tone of Wording</i> Smoking in public	Allow	Forbid
<i>Open vs. Closed Form on Work Values</i>	Closed	Open

as quickly as possible, reinforcing the tendency of respondents to give the first answer that comes to mind, whereas in the self-administered survey respondents had been encouraged to take their time in answering.¹⁰

Each of these procedures was designed to strengthen the response effects of the manipulation of the question forms, making the differences between the two modes of data collection large enough to detect statistically with small samples and categorical variables.

Only eight of 163 subjects assigned to the telephone condition could not be contacted, resulting in a response rate of 95.1 percent. A total of 194 self-administered questionnaires and 155 telephone interviews were completed. From the self-administered group 11 people who reported not having a telephone were eliminated from the analysis to make the two samples more comparable. This resulted in a total of 183 subjects in the self-administered condition, a relatively small number as compared to that in the U.S. experiment.

The replication in West Germany, then, was more like a typical laboratory experiment, conducted partly under field conditions, whereas the U.S. experiment was more of a field study similar to the usual survey. These variations in implementing the experiment, however, make it all the more valuable if the results should replicate from one setting to the other.

2. FINDINGS

2.1 Response Order Effects

Divorce Issue. Schuman and Presser (1981, Chap. 2) have discovered that when respondents are asked whether a divorce should "be *easier to obtain, more difficult to obtain, or stay as it is now,*" they are significantly more likely to select the middle response alternative, *stay as it is now,* when it is offered in the last, rather than the second (or middle) position in the response sequence. We had hypothesized that response order would make a significant difference in the results for the divorce item in our telephone surveys, as it did in Schuman and Presser's studies, but that it would have no significant effect on the results of the self-administered surveys. Surprisingly, however, we found that the order in which the middle alternative was presented in the divorce item made no significant

¹⁰ In the self-administered survey the manipulation was strengthened, in both experiments, by the instructions to "...take your time and read each question carefully before answering the questionnaire."

difference in the results, in either the telephone survey or the self-administered survey, in either the U.S. or the German experiment (data not shown here)¹¹. This response order effect may not replicate with college students because the issue of divorce does not have the same psychological significance for them as it does for older adults, many of whom have had to make a decision about divorce in their lives, or who are contemplating such a decision. Probing respondents who select the answer, "stay as it is now," with a followup question about why they chose that alternative, might reveal differences in the meaning of divorce for these two populations.

Nuclear Power Issue. When respondents are asked, "Are you in favor of building more nuclear power plants, would you prefer to see all nuclear power plants closed down, or would you favor operating only those that are already built?," they are significantly more likely to choose the middle alternative, *operating only those that are already built*, when it is presented in the last rather than the second (or middle) position in the response sequence (Bishop, 1987). As with the divorce item, we had hypothesized that response order would make a significant difference in the results for the nuclear power item in our telephone surveys, and little or no difference in the findings from the self-administered surveys. The figures in Table 2 tend to confirm this hypothesis. When interviewed by telephone, respondents in the U.S. experiment were noticeably more likely to select the middle alternative if it was presented in the last, rather than the middle, position. The German data showed the same pattern: the difference between the two question forms in the telephone survey was sizable (15 percent) and close to being statistically significant, despite the small subsample sizes. When respondents were given the self-administered form, however, the order in which the middle alternative was presented in the nuclear power question made no significant difference in the percentage of respondents who chose it in either of the two experiments. Though there was a tendency for German respondents to select the middle alternative more often when it was presented in the middle rather than the last position in the self-administered survey, this pattern is probably due to chance since it did not replicate in the U.S. experiment.

¹¹ To conserve space, this and several other tables of data have been omitted from this chapter, but are available from the authors on request.

Table 2. Response to Nuclear Power Plants Item by Question Form and Mode of Data Collection

	Telephone		Self-Administered	
	Middle Alt. Offered in Last Position	Middle Alt. Offered in Middle Position	Middle Alt. Offered in Last Position	Middle Alt. Offered in Middle Position
Nuclear plants (U.S.)				
Build more	31.9%	32.6%	32.5%	38.7%
Close all	17.8	26.7	21.7	18.8
Operate only those built	50.3	40.6	45.8	42.5
	100.0 (185)	99.9 (187)	100.0 (166)	100.0 (181)
	X^2 (middle vs. other responses) = 3.10, $df=1$, $p=.078$ X^2 (middle vs. other responses) = 0.25, $df=1$, $p>.25$ Three-way interaction (response by form by mode): $X^2=0.74$, $df=1$, n.s.			
Nuclear plants (Germany)				
Build more	14.5%	18.7%	25.3%	20.9%
Close all	15.8	26.7	23.1	15.4
Operate only those built	69.7	54.7	51.6	63.7
	100.0 (76)	100.1 (75)	100.0 (91)	100.0 (91)
	X^2 (middle vs. other responses) = 3.67, $df=1$, $p=.055$ X^2 (middle vs. other responses) = 2.73, $df=1$, $p<.098$ Three-way interaction (response by form by mode): $X^2=6.40$, $df=1$, $p<.05$			

2.2 Question Order

Japanese Trade Issue. As expected, when respondents in the U.S. experiment were interviewed by telephone, the order in which the questions about Japanese-American trade relations were asked made a sizable and significant difference in the results (see Table 3). Indeed, the results were remarkably similar to those reported by Schuman and Ludwig (1983): respondents were significantly more likely to favor limiting Japanese imports to the United States (69.4 percent) than they were to favor limiting U.S. exports to Japan (53.8 percent) when each question was asked in the first position ($\chi^2 = 9.6$, $df = 1$, $p < .01$). Similarly, we found that support for limiting U.S. exports to Japan (67.9 percent) increased significantly when respondents were asked about it immediately *after* the question about trade restrictions on Japanese imports by the United States, presumably because a norm of even handedness had been evoked by the sequence of the questions. Furthermore, we found, as did Schuman and Ludwig, that the norm of even handedness does not necessarily operate with equal force in both directions: support for limiting Japanese imports (67.2 percent) did not decline significantly, if at all, when respondents were asked about it immediately *after* the question about limiting U.S. exports to Japan. This asymmetry, as Schuman and Ludwig suggest, may be the result of American perceptions that the "unfair" Japanese competition for the U.S. market is what needs to be righted by restrictions on imports.

But when respondents were asked these same questions in the self-administered form, the order in which they were presented had, as predicted, no significant effect on the results. This does not mean, however, that the norm of even handedness had no influence on responses to the questions about Japanese-American trade relations in the self-administered form. To the contrary, because respondents were able to look at both of the questions about trade restrictions simultaneously, they could not help but realize that a norm of even handedness was called for in answering the questions. And that is why we find, unlike the results of the telephone survey, that respondents were *not* significantly more likely to favor trade restrictions by the United States (68.1 percent) than they were to favor restrictions by Japan (64.5 percent) when each question was asked in the first position ($\chi^2 = 0.53$, $df = 1$, n.s.). Indeed, the absence of an order effect on the responses to these questions in the self-administered form is precisely what Schuman and Presser's hypothesis would predict.

The same question order effect occurred in the German experiment, even though the marginal distributions of responses to the trade items were quite different from those in the United States (i.e., more favorable

Table 3. Response to Japanese Trade Items by Question Form and Mode of Data Collection

	Telephone		Self-Administered	
	Limit U.S. Item Asked Before Limit Japan Item	Limit U.S. Item Asked After Limit Japan Item	Limit U.S. Item Asked Before Limit Japan Item	Limit U.S. Item Asked After Limit Japan Item
Should Japan limit U.S. imports				
Yes	53.8%	67.9%	64.5%	67.6%
No	46.2	32.1	35.5	32.4
	100.0 (186)	100.0 (187)	100.0 (166)	100.0 (182)
	$\chi^2=7.26, df=1, p<.01$		$\chi^2=0.25, df=1, p>.25$	
	Three-way interaction (response by form by mode): $\chi^2=2.17, df=1, .10<p<.25$			
Should U.S. limit Japanese imports				
Yes	67.2%	69.4%	68.1%	68.1%
No	32.8	30.6	31.9	31.9
	100.0 (186)	100.0 (186)	100.0 (166)	100.0 (182)
	$\chi^2=0.11, df=1, p>.25$		$\chi^2=0.00, df=1, p>.25$	
	Three-way interaction (response by form by mode): $\chi^2=0.09, df=1, p>.25$			
	Limit Germany Item Asked Before Limit Japan Item	Limit Germany Item Asked After Limit Japan Item	Limit Germany Item Asked Before Limit Japan Item	Limit Germany Item Asked After Limit Japan Item
Should Japan limit German imports				
Yes	12.8%	30.7%	30.0%	25.0%
No	87.2	69.3	70.0	75.0
	100.0 (78)	100.0 (75)	100.0 (90)	100.0 (92)
	$\chi^2=7.34, df=1, p<.01$		$\chi^2=0.57, df=1, p>.25$	
	Three-way interaction (response by form by mode): $\chi^2=6.58, df=1, p<.02$			
Should Germany limit Japanese imports				
Yes	24.4%	36.5%	41.1%	33.7%
No	75.6	63.5	58.9	66.3
	100.0 (78)	100.0 (74)	100.0 (90)	100.0 (92)
	$\chi^2=2.65, df=1, p=.103$		$\chi^2=1.07, df=1, p>.25$	
	Three-way interaction (response by form by mode): $\chi^2=3.64, df=1, p=.056$			

toward "free trade"). When respondents were interviewed by telephone, the sequence of the questions made a significant difference in the results, but it made little or no difference when the respondents were given a self-administered questionnaire. Unlike the results of the U.S. experiment, however, a norm of even handedness appears to have influenced responses to both questions more equally. Support for limiting Japanese imports to Germany declined (24.4 percent) when respondents were asked about it immediately *after* the question about trade restrictions by Japan, though the difference was not statistically significant. The effect of the norm may be more symmetrical in the German case because trade relations with Japan are probably not viewed as unbalanced as they are in the United States. In other words, a necessary condition for the norm to operate with equal force in both directions may be a perception that both parties (e.g., nations) are presently engaged in fair and equal competition. Otherwise, the effect of the norm will be asymmetrical, acting to equalize the "unfair" competition, as in the U.S. case.

Abortion Issue. Here too, as hypothesized, the order of the questions made a difference in the results, but only when respondents were interviewed by telephone (cf. Schuman and Presser, 1981; Bishop et al., 1985). In both experiments, respondents were more likely to approve of an abortion for a woman who is married and does not want any more children when they were asked about it on the telephone, and *before* the question about abortion in the case of a possible birth defect, than when they were asked about it *after* the latter question (see Table 4). In the self-administered questionnaire, however, the sequence of the questions made little or no difference in the results. Though the evidence for the hypothesis was statistically significant only in the German experiment, the pattern in the two studies was sufficiently similar that it is highly unlikely to be the result of chance.

2.3 Middle Response Alternatives

As in previous studies (Bishop, 1987; Schuman and Presser, 1981, Chap. 6) we found that respondents were much more likely to select a middle response alternative if it was explicitly offered to them than if it was not (data not shown here). This pattern occurred on both the question about marijuana penalties and the item on left/right political identification, and in both experiments. And, as predicted, this question form effect was just as likely to occur in a self-administered survey as in a telephone survey, and to about the same degree in each experiment.

We also discovered that respondents were significantly more likely to select the middle response alternative on Form A of the question about

Table 4. Response to Abortion Item by Question Form and Mode of Data Collection

	Telephone		Self-Administered	
	Women's Right Item Asked After Birth Defect Item	Women's Right Item Asked Before Birth Defect Item	Women's Right Item Asked After Birth Defect Item	Women's Right Item Asked Before Birth Defect Item
Abortion if woman does not want any more children (U.S.)				
Yes (allow)	51.9%	59.3%	51.2%	47.5%
No	48.1	40.7	48.8	52.6
	100.0	100.0	100.0	100.0
	(185)	(189)	(166)	(181)
	$\chi^2=1.77, df=1, .10 < p < .25$		$\chi^2 = .34, df=1, p > .25$	
	Three-way interaction (response by form by mode): $\chi^2=2.22, df=1, .10 < p < .25$			
Abortion if woman does not want any more children (Germany)				
Yes (allow)	42.1%	69.3%	49.5%	58.2%
No	57.9	30.7	50.5	41.8
	100.0	100.0	100.0	100.0
	(76)	(75)	(91)	(91)
	$\chi^2=11.50, df=1, p < .01$		$\chi^2=1.42, df=1, .10 < p < .25$	
	Three-way interaction (response by form by mode): $\chi^2=2.98, df=1, p=.084$			

defense spending than the no opinion alternative on Form B of the question (data not shown here), clearly indicating that these two response alternatives are not psychologically equivalent. As would be expected, however, this question form effect was observed in both the telephone survey and the self-administered survey, and in both experiments. In other words, this form effect, just like those in the previous experiments with the marijuana and left/right items, did not interact significantly with the mode of data collection in either of the two experiments.

2.4 The No Opinion Alternative

Not surprisingly, as with middle responses, both experiments showed that respondents were significantly more likely to choose a no opinion alternative if it was explicitly offered to them than if it was not, and not only on the real issue of Arab-Israeli relations, but also on the fictitious "International Trade Act of 1986" (data not shown). Again, as expected,

this form effect was just as evident in the self-administered survey as it was in the telephone survey, and in both investigations. Question form effects, then, whether the result of the presence or absence of a no opinion alternative or a middle response alternative, do not appear to depend upon the mode of data collection, though we have obviously not tested this hypothesis fully through comparisons with data from face to face interviews (cf. this volume, Chapters 18, 19, and 21).

2.5 Tone of Wording

The figures in Table 5 on the forbid/allow effect are difficult to explain. As Schuman and Presser (1981) would predict, the effect does not generalize very well, if at all, to a concrete subject such as regulating smoking in public places. But their prediction seems to apply only when the data are collected by telephone. In both experiments, we found no significant differences by question form in the telephone survey. The data for the self-administered surveys, however, are much harder to interpret. In the U.S. experiment we discovered, contrary to the results of all previous research on the forbid/allow effect, that respondents were significantly more likely to say that something, such as smoking in public places, should be *forbidden* than they were to say that it should *not be allowed*, whereas in the German experiment the results were exactly the opposite: respondents were much more likely to say that smoking in public places should *not be allowed* than they were to say it should be *forbidden*, the pattern we would expect to have found, if any. Suffice it to say that a further replication is in order.

2.6 Open vs. Closed Question Form

A comparison of the responses to the open and closed form of the work values question in Table 6 shows the following:

1. As in previous experiments by Schuman and Presser (1981, Chap. 3), most responses to the closed form of the question fell within the first five precoded categories, whereas responses to the open form spread much more widely beyond these five categories. This pattern occurred in both the self-administered and the telephone surveys, and in both experiments, though it was somewhat sharper in the U.S. data than in the German data. For whatever reason, the German students were significantly more likely than U.S. students to volunteer more than one response to the question on both the open and closed forms, and in both

Table 5. Response to Forbid-Allow Items by Question Form and Mode of Data Collection

	Telephone		Self-Administered	
	Allow Form	Forbid Form	Allow Form	Forbid Form
Smoking in public places (U.S.)				
Yes (allowed, not forbidden)	51.1%	47.6%	53.0%	38.7%
No (not allowed, forbidden)	48.9	52.4	47.0	61.3
	100.0 (180)	100.0 (189)	100.0 (166)	100.0 (181)
	$\chi^2=0.32, df=1, p>.25$		$\chi^2=8.61, df=1, p<.02$	
	Three-way interaction (response by form by mode): $\chi^2=2.18, df=1, .10<p<.25$			
Smoking in public places (Germany)				
Yes (allowed, not forbidden)	66.7%	72.4%	47.1%	67.4%
No (not allowed, forbidden)	33.3	27.6	52.9	32.6
	100.0 (78)	100.0 (76)	100.0 (86)	100.0 (92)
	$\chi^2=0.59, df=1, p>.25$		$\chi^2=7.53, df=1, p<.01$	
	Three-way interaction (response by form by mode): $\chi^2=1.50, df=1, .10<p<.25$			

the self-administered and the telephone surveys. Notice too that respondents in both studies who received the open form of the question in the self-administered survey were more likely to have given more than one response to the question (by writing it in) than those who received the open form of the question in the telephone survey, most likely because of the inability to probe and clarify such responses in the self-administered condition.

2. If we examine the data for just the first five categories common to both forms, it appears that there was a substantial difference between the open and closed forms in the percentage choosing a *feeling of accomplishment* as the most important work value. As Schuman and Presser have discovered, respondents were much more likely to select the *feeling of accomplishment* category when it was explicitly offered to them on the closed form than to volunteer it on the open form. This response pattern was evident in both the telephone and self-administered surveys in the U.S. sample, but only in the telephone survey in the German sample. Some of the difference between the open and closed form on the Accomplishment category, however, may be due, as Schuman and Presser

have suggested, to the fact that many respondents who are coded into the Satisfaction category on the open form (see Table 6) would, if properly probed, end up in the Accomplishment category. Many of the respondents who gave more than one codable response, particularly on the open form of the self-administered survey, might also have selected Accomplishment as the most important work value if forced to choose. So the apparent difference between the two forms may represent primarily variations in coding and probing procedures on the open form of the question. Other differences between the two forms on the common categories seem to be relatively minor, especially given the small subsample sizes on which they are based, especially in the German study.

3. On the *closed* form of the question respondents in the U.S. sample were significantly more likely to select one of the last three of the five common response alternatives (control of work, pleasant work, security) when they were interviewed on the telephone than when they were given the self-administered form ($\chi^2 = 4.27$, $df = 1$, $p < .04$). But this response order effect did not replicate in the German sample. So the results are somewhat ambiguous, pending a further replication, preferably with a much larger sample.

3. CONCLUSION

Though the results of this crosscultural experiment and replication were not as unequivocal as we might have liked, they clearly suggest that question order and response order effects are significantly less likely to occur in a self-administered survey than in a telephone survey, whereas question form and wording effects are probably just as likely to occur with one mode of data collection as another. To the extent that such response effects are regarded as unwanted systematic sources of error in survey measurement, our findings on question order and response order effects would indicate that the quality of data gathered through self-administered surveys may, other things being equal (e.g., response rate, respondent literacy), be better than that obtained by telephone surveys [see this volume, Chapters 18, 19, and 21 for comparisons of data quality in telephone and face to face interviews]. Further replications of these findings with similar, as well as different, populations would certainly be useful, as would extensions to other topics, response effects, and variations in modes of data collection (face to face interviews and standard mail surveys). For we now know that generalizations about response effects in surveys are even more conditional than we thought they were.

Table 6. Response to Work Values Item by Question Form and Mode of Data Collection

	Telephone		Self-Administered	
	Open	Closed	Open	Closed
Most prefer in a job (U.S)				
Pays well	8.4%	8.6%	6.7%	14.5%
Feeling of accomplishment	16.3	49.5	22.2	53.6
Control of work	8.4	16.1	3.3	12.0
Pleasant work	11.1	21.5	8.9	14.5
Security	8.4	4.3	1.1	3.6
Liking/satisfaction	30.5	0.0	8.3	0.0
Promotion opportunity	3.2	0.0	5.6	0.0
More than one response	3.7	0.0	41.1	1.2
Other	10.0	0.0	2.8	0.6
	100.0	100.0	100.0	100.0
	(190)	(186)	(180)	(166)
Most Prefer in a Job (Germany)				
Pays well	0.0%	1.3%	1.1%	2.2%
Feeling of accomplishment	15.5	42.1	15.6	11.1
Control of work	2.8	21.1	3.3	22.2
Pleasant work	11.3	14.5	2.2	23.3
Security	1.4	7.9	0.0	3.3
Liking/satisfaction	19.7	0.0	11.1	0.0
Promotion opportunity	1.4	0.0	1.1	0.0
More than one response	42.3	6.6	56.7	36.7
Other	5.6	6.6	8.8	1.1
	100.0	100.0	100.0	100.0
	(71)	(76)	(90)	(90)

APPENDIX. WORDING OF THE QUESTIONS IN THE SELF-ADMINISTERED SURVEY

Form A

Form B

In your opinion, should divorce in this country be...

1. easier to obtain
2. more difficult to obtain
3. stay as it is now

Do you think that the Japanese government should be allowed to set limits on how much American industry can sell in Japan?

Do you think that the American government should be allowed to set limits on how much Japanese industry can sell in the United States?

1. Yes
2. No

Do you think it should be possible for a pregnant woman to obtain a legal abortion if there is a strong chance of serious defect in the baby?

1. Yes
2. No

Do you think it should be possible for a pregnant woman to obtain a legal abortion if she is married and does not want any more children?

1. Yes
2. No

Do you think that smoking in public places, such as restaurants, should be allowed?

1. Yes
2. No

1. In your opinion, should divorce in this country be...

1. easier to obtain
2. stay as it is now
3. more difficult to obtain

2. Do you think that the American government should be allowed to set limits on how much Japanese industry can sell in the United States?

3. Do you think that the Japanese government should be allowed to set limits on how much American industry can sell in Japan?

1. Yes
2. No

4. Do you think it should be possible for a pregnant woman to obtain a legal abortion if she is married and does not want any more children?

1. Yes
2. No

5. Do you think it should be possible for a pregnant woman to obtain a legal abortion if there is a strong chance of serious defect in the baby?

1. Yes
2. No

6. Do you think that smoking in public places, such as restaurants, should be forbidden?

1. Yes
2. No

Form A

Form B

7. Some people say that the United States needs to develop new (alternative) power sources from nuclear energy in order to meet our needs for the future. Other people say that the danger to the environment and the possibility of accidents are too great. What do you think—do you...

1. favor building more nuclear power plants
2. prefer to see all nuclear power plants closed down
3. favor operating only those that are already built

8. In your opinion, should penalties for using marijuana be...

1. more strict
2. less strict

9. Some people believe we should spend less money for defense. Others feel that defense spending should be increased. How about you—do you think defense spending should be...

1. increased
2. decreased
3. continued at the present level

10. The United Nations has been considering the International Trade Act of 1986. Do you...

1. favor the passage of this act
2. oppose the passage of this act

11. On most political issues, would you say you are on the left, on the right, or in the middle?

1. Left
2. Right
3. Middle

7. Some people say that the United States needs to develop new (alternative) power sources from nuclear energy in order to meet our needs for the future. Other people say that the danger to the environment and the possibility of accidents are too great.

What do you think—do you...

1. favor building more nuclear power plants
2. favor operating only those that are already built
3. prefer to see all nuclear power plants closed down

8. In your opinion, should penalties for using marijuana be...

1. more strict
2. less strict
3. about the same as they are now

9. Some people believe we should spend less money for defense. Others feel that defense spending should be increased. How about you—do you think defense spending should be...

1. increased
2. decreased
3. no opinion

10. The United Nations has been considering the International Trade Act of 1986. Do you...

1. favor the passage of this act
2. oppose the passage of this act
3. no opinion

11. On most political issues, would you say you are on the left or on the right?

1. Left
2. Right

Form A

12. Do you think the Arab Nations are trying to defeat Israel, trying to work for a real peace with Israel, or do you not have an opinion on that?

1. Trying to defeat Israel
2. Trying to work for a real peace with Israel
3. No opinion

13. [Question about the Falkland Islands]

14. [Question about Afghanistan]

15. The next question is on the subject of work. People look for different things in a job. Which *one* of the following five things would you most prefer in a job...

1. work that pays well
2. work that gives a feeling of accomplishment
3. work where there is not too much supervision and you make most decisions yourself
4. work that is pleasant and where the other people are nice to work with
5. work that is steady with little chance of being laid off

Form B

12. Do you think the Arab Nations are trying to defeat Israel, or are they trying to work for a real peace with Israel?

1. Trying to defeat Israel
2. Trying to work for a real peace with Israel

13. [Question about the Falkland Islands]

14. [Question about Afghanistan]

15. The next question is on the subject of work. People look for different things in a job. What would you *most prefer* in a job?

CHAPTER 21

THE QUALITY OF INCOME INFORMATION IN TELEPHONE AND FACE TO FACE SURVEYS

Eszter Körmendi¹

The Danish National Institute of Social Research

1. PROBLEMS OF ASKING INCOME QUESTIONS IN SURVEYS

There appears to be a great deal of agreement in Western industrialized countries on the type of questions respondents perceive as being threatening or embarrassing.

"Threatening questions" encompass questions where the answers may cause the respondents to fear a lowering of their esteem in the eyes of others, i.e., questions relating to social desirability.

Such questions can either involve activities which are considered embarrassing, strictly private, illegal, etc., or conversely, activities regarded as desirable, leading to higher social status, being "with it," etc. A classic example, and the type of question most often treated in textbooks, is income questions (Sudman and Bradburn, 1982). Difficulties can arise from both cognitive and emotional circumstances.

First, a memory factor is involved, especially when the question concerns income earned the previous year. This difficulty is somewhat lesser in face to face interviews by means of a show card with relatively large income brackets which help the interviewee to find the correct bracket. Furthermore, the large intervals on the card may reassure the interviewee that only an approximate rather than an exact knowledge is required.

¹The author thanks soc.drs. J. Noordhoek, Danmarks Statistik, for his comments on an earlier version of this paper and for his practical help that made the comparison between the self-reported and register data possible.