The impact of administration mode on response effects in survey measurement

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

Veröffentlichungsversion / Published Version
Arbeitspapier / working paper

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:
Schwarz, N., Strack, F., Hippler, H.-J., & Bishop, G. (1990). *The impact of administration mode on response effects in survey measurement.* (ZUMA-Arbeitsbericht, 1990/02). Mannheim: Zentrum für Umfragen, Methoden und Analysen - ZUMA-. https://nbn-resolving.org/urn:nbn:de:0168-ssoar-67199

The Impact of Administration Mode on Response Effects in Survey Measurement

Norbert Schwarz, Fritz Strack
Hans-J. Hippler, George Bishop

ZUMA-Arbeitsbericht Nr. 90/02

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. 90/02 by the same authors.

Norbert Schwarz, Fritz Strack, Hans-J. Hippler, & George Bishop (1991). The impact of administration mode on response effects in survey measurement. In J. Jobe & E. Loftus (Eds.), Cognitive aspects of survey methodology. Special issue of:
Applied Cognitive Psychology, 1991, 5, 193-212.

APPLIED COGNITIVE PSYCHOLOGY, VOL. 5, 193–212 (1991)

The Impact of Administration Mode on Response Effects in Survey Measurement

NORBERT SCHWARZ

Zentrum für Umfragen, Methoden und Analysen, ZUMA, Mannheim, F.R.G.

FRITZ STRACK

Universität Mannheim, Mannheim, F.R.G.

HANS-J. HIPPLER

Zentrum für Umfragen, Methoden und Analysen, ZUMA, Mannheim, F.R.G.

and GEORGE BISHOP

Institute for Policy Research, University of Cincinnati, Cincinnati, Ohio, U.S.A.

SUMMARY

The major differences between face-to-face and telephone interviews as well as self-administered questionnaires are reviewed and are related to the cognitive and communicative processes assumed to underlie the process of question answering. Based on these considerations the impact of administration mode on the emergence of well-known response effects in survey measurement is discussed, and relevant experimental evidence is reported. It is concluded that administration mode affects the emergence of question order and context effects; the emergence of response order effects; the validity of retrospective reports; and the degree of socially desirable responding. The emergence of question wording and question form effects, on the other hand, appears to be relatively independent of administration mode.

That the results of public opinion surveys can be significantly affected by the way in which questions are worded, the form in which they are presented, and the order or context in which they are asked is well known. While a considerable number of these influences have been documented in the literature (cf. Dijkstra and van der Zouwen, 1982; Payne, 1951; Schuman and Presser, 1981; Sudman and Bradburn, 1974 for reviews), the underlying cognitive processes have only recently received systematic attention (cf. Hippler, Schwarz, and Sudman, 1987; Jabine, Straf, Tanur, and Tourangeau, 1984; Schwarz and Sudman, in press for examples). Not surprisingly, all researchers agree that answering a survey question requires that respondents solve several tasks (see Strack and Martin, 1987; Tourangeau, 1984, 1987; Tourangeau and Rasinski, 1988 for detailed discussions). As a first step, respondents have to interpret the question to understand what is meant. If the question is an opinion question, they subsequently have to 'generate' an opinion on the issue. To do so, they need to retrieve relevant information from memory to form a judgement. Alterna-

0888–4080/91/030193–20$10.00

tively, they may retrieve a previously formed judgement on the issue, if accessible in memory. If the question asks them to report on a certain behaviour or personal experience, on the other hand, they have to retrieve relevant instances of that behaviour from memory. Depending on the nature of the question they may also need to determine the frequency of this behaviour or the date of its occurrence (see Bradburn, Rips, and Shevell, 1987; Schwarz, 1990 for reviews of these latter tasks). Once a 'private' judgement is formed in respondents' minds, they have to communicate it to the interviewer. To do so they may need to format their judgement to fit the response alternatives provided by the researcher. Moreover, respondents may wish to edit their response before they communicate it, due to influences of social desirability and situational adequacy.

Accordingly, interpreting the question, generating an opinion, formating the response, and editing are the main psychological components of a process that starts with respondents' exposure to a survey question and ends with their overt response. Detailed discussions of these steps have been provided elsewhere (Strack and Martin, 1987; Tourangeau and Rasinski, 1988), and do not need reiteration. However, each of these operations may be affected by psychological variables that are likely to covary with the mode of data collection, and this possibility is of primary interest in the present paper. Specifically, we will review the major psychological differences between face-to-face and telephone interviews as well as self-administered questionnaires, elaborating on their potential relevance to the cognitive processes that are assumed to underlie the process of question answering. Where available, we will report experimental evidence that bears on the impact of administration mode on the major types of response effects in survey measurement, and where this evidence is missing we will point out some of the research issues that need to be addressed in future studies.

MODES OF DATA COLLECTION AND THE PROCESS OF QUESTION ANSWERING

Table 1 shows a summary of the key differences between face-to-face and telephone interviews as well as self-administered questionnaires.

Visual vs. auditory presentation of the stimuli

One of the most obvious differences between the modes of administration is the sensory channel in which the material is presented. In self-administered questionnaires the items are visually displayed to the respondent who has to read the material. In telephone interviews, at the other extreme, the items and the response alternatives are read to respondents who have to listen to what the interviewer says. In face-to-face interviews both modes of presentation may occur.

Sequential vs. simultaneous presentation of the items

Closely related to the previous distinction is the temporal order in which the material is presented. Telephone and face-to-face interviews have a strict sequential organization. That is, respondents have to process the information in the temporal succession

Table 1. Comparison of psychological aspects of mode of survey data collection

Variable	Face-to-face interview	Telephone interview	Self-administered questionnaire
Visual (V) vs. auditory (A) presentation	A/V	A	V
Sequential (SE) vs. simultaneous (SI) presentation	SE	SE	SI
Time pressure (+/−)	+	++	0
Additional explanations from interviewer (+/−)	++	+	0
Perception of interviewer characteristics (+/−)	++	+	0
Perceived confidentiality (+/−)	−−	−	+/?
External distractions	?	?	?

and the pace in which it is presented by the interviewer. They usually cannot go back and forth or spend relatively more or less time on some particular item. And even if respondents are allowed to return to previous items should they want to correct their responses, they rarely do so, in part because tracking one's previous responses presents a difficult memory task under telephone and face-to-face conditions. In contrast, keeping track of one's responses, and going back and forth between items, pose no difficulties under self-administered questionnaire conditions. Here respondents can use as much time as they want to work on the questionnaire. Even if the questionnaire is administered in a classroom setting, in which the available amount of time is limited, they can at least allocate the time provided to them to those questions that they want to think about more carefully. Moreover, a self-administered questionnaire allows respondents to go back to previous questions and to be reminded of their earlier answers. At the extreme, respondents may complete different parts of the questionnaire at different times. Accordingly, we may expect that self-administered questionnaires render the sequential organization of questions less influential.

Time pressure

Time pressure is a psychologically relevant variable that has been shown to increase 'top of the head' phenomena. Most importantly, time pressure interferes with extensive recall processes and increases reliance on the first thing that comes to mind (see Bodenhausen and Wyer, 1987; Kruglanski, 1980). Moreover, it induces individuals to resort to heuristic processing strategies (Kahneman, Slovic, and Tversky, 1982) at the expense of detail-oriented piecemeal processing strategies (e.g. Strack, Erber, and Wicklund, 1983). Accordingly, time pressure is likely to affect recall as well as judgemental processes, as will be elaborated below.

The greatest time pressure can be expected under telephone interview conditions, where moments of silent reflection cannot be bridged by non-verbal communication that indicates that the respondent is still paying attention to the task (Ball, 1968; Groves and Kahn, 1979). The least degree of time pressure is induced by self-administered questionnaires that allow respondents to work at their own pace. Face-to-face

interviews create intermediate time pressure, due to the possibility of bridging pauses by non-verbal communication.

Interviewer–respondent interaction

While social interaction is severely constrained in all standardized survey interviews, the modes of data collection clearly differ in the degree to which they restrict non-verbal communication. While face-to-face interviews provide full access to the non-verbal cues of the participants, participants in telephone interviews are restricted to paraverbal cues, whereas social interaction is largely absent under self-administered conditions.

Psychological research has identified various functions of non-verbal cues during face-to-face interaction (see Argyle, 1969 for a review). Most importantly, non-verbal cues serve to indicate mutual attention and responsiveness, provide feedback as well as illustrations for what is being said (in the form of gestures), and convey interpersonal attitudes. Although laboratory research on telephone and face-to-face interaction in problem-solving situations suggests that the absence of visual contact may have only small and elusive effects on information transmission (Reid, 1977; Williams, 1977), the degree of mutual contact does affect respondents' opportunity to receive additional explanations from the interviewer, as well as the likelihood of interviewer effects that are based on respondents' perceptions of interviewer characteristics. We will consider each of these possibilities in turn.

Additional explanations from the interviewer

In face-to-face interviews, where the interviewer can monitor the respondent's non-verbal expressions, and to a lesser degree under telephone interview conditions, where the interviewer is limited to monitoring the respondent's verbal utterances, respondents may be given additional information by the interviewer. Under both of these conditions they are free to request additional information should they desire to do so. Even though the additional information is usually restricted to certain prescribed feedback, it may help the respondent to determine the meaning of the questions. In fact even the uninformative—but not unusual—clarification, 'whatever it means to you', may be likely to short-cut further attempts of the respondent to screen question context in search for an appropriate interpretation. Under self-administered questionnaire conditions, on the other hand, the respondent is much more dependent on the context that is explicitly provided by the questionnaire to draw inferences about the intended meaning of the questions (cf. Schwarz and Strack, 1988; Strack and Martin, 1987)—and has the time and opportunity to consider related questions to disambiguate the meaning of obscure items.

Perception of interviewer characteristics

Interviewer characteristics are more likely to be noticed by the respondent when he or she has face-to-face contact than when the interviewer cannot be seen, as is the case under telephone interview conditions, where the identification of interviewer characteristics is limited to characteristics that may be inferred from paralinguistic cues and speech styles (such as sex, age, or race). Under both conditions,

however, interviewers may (unconsciously) convey their personal attitudes or their (dis)approval to the respondent, although this seems the more likely the more the respondent can monitor the interviewer's expression.

Under self-administered questionnaire conditions, of course, no interviewer is required, although respondents may pick up characteristics of the researcher from the cover letter, the person who dropped off the questionnaire, and so on. While respondents' perception of interviewer characteristics may increase socially desirable responses, it may also serve to increase rapport with the interviewer, rendering the potential impact of this variable ambivalent.

Perceived confidentiality

Survey responses may be more or less confidential with regard to the interviewer or researcher, as well as other household members. At least in principle, self-administered questionnaires that may be returned without identifying information provide the highest degree of confidentiality *vis-à-vis* the researcher. Under face-to-face conditions, and to a lesser degree under telephone interview conditions, however, the respondent is known at least to the interviewer, which may increase socially desirable responding. Confidentiality *vis-à-vis* other household members, on the other hand, may be best reinforced under face-to-face interview conditions and is completely left to the respondent under self-administered conditions. At the extreme, the respondent may involve others in answering the questions.

External distractions

While external distractions, e.g. due to the presence of children or other household members, cannot be excluded under any administration mode, they can be monitored by the interviewer under face-to-face conditions and to some degree under telephone interview conditions. Most importantly, general rules of politeness suggest that ongoing conversations are not to be disrupted. External distractions may be more likely under mail survey conditions, where the questionnaire may be completed while watching TV, or the like. On the other hand, respondents are free to work on a self-administered questionnaire at a time of their choice, rendering the relationship between mode and external distractions ambiguous.

Differential self-selection of respondents

Finally, different administration modes may result in differential self-selection of respondents with different characteristics. In general, respondents with a low level of education are assumed to be underrepresented in mail surveys relative to face-to-face and telephone interviews (e.g. Dillman, 1978). Moreover, respondents in mail surveys are assumed to be more interested in the topic of the survey because they can preview the questions before they decide to participate (Dillman, 1978). Accordingly, respondents' cognitive sophistication, as well as their motivation, may vary across modes. However, this possibility is not yet well documented.

Given that different modes of data collection do result in markedly different response rates (see Groves and Lyberg, 1988), differential self-selection of respondents along any number of variables provides a plausible account for many of the observed

mode effects. Accordingly, the use of comparable samples with high response rates and tight controls for self-selection processes is required to make mode comparisons meaningful (see Bishop, Schwarz, Hippler, and Strack, 1988).

IMPLICATIONS FOR RESPONSE EFFECTS

We will now turn to the implications of the above differences between modes of data collection for the cognitive and communicative processes that are assumed to underlie different response effects, reporting experimental evidence where it is available. As will soon become evident, adequately controlled empirical studies are rare. This state of affairs is primarily due to the applied interest that governs research on mode effects in survey measurement. As Biemer (1988, p. 274) points out, the objective of mode comparisons in the view of survey practitioners is 'to compare the quality of data from an efficiently designed face-to-face (or mail) survey to an equally well designed telephone survey'. Accordingly, 'the mode of interview will typically *not* be the only design factor that will differ'. Rather, 'the comparison is really between two *systems of data collection*, each with design parameters and procedures that may be broadly equivalent, yet particularly adapted for efficiency in the given mode of interviewing'. As a result the exact source of any observed differences can usually not be identified—and the available studies are often not intended to do so. Thus, appropriately focused experimental tests are often not available and some of the subsequently offered considerations are only weakly supported, although highly plausible on the basis of current theorizing. What is urgently needed are tightly controlled studies that allow the identification of the *sources* of differential response effects under different administration modes. We hope that the present selective review of the currently available evidence will stimulate future experimental work in this theoretically as well as practically important area.

In the following sections we will first review response effects that are likely to vary as a function of administration mode, and will subsequently address response effects that appear to be less affected by administration mode.

Question order and question context

Effects of the sheer *order* in which questions are asked require sequential question presentation. Most question order effects should therefore be either reduced or absent under self-administered questionnaire conditions, depending on the proportion of respondents who read all or some of the questions before answering them, thus eliminating sequential presentation.

However, the absence of question *order* effects does not imply that the broader *context* of a question is generally unlikely to affect responses under self-administered conditions. Rather, it only implies that the impact of question context should be less dependent on the *order* in which the questions are asked. In fact, some context effects may be *more* likely under self-administered conditions than under face-to-face or telephone interview conditions, whereas other context effects may be *less* likely.

On the one hand, preceding questions have been shown to increase the cognitive accessibility of information that is used in answering them. This increases the likelihood that this information will be considered in answering subsequent, related ques-

tions because individuals do not retrieve all information that potentially bears on a judgement, but truncate the search process as soon as enough information has come to mind to form a judgement (see Bodenhausen and Wyer, 1987; Strack, Martin, and Schwarz, 1988; Tourangeau and Rasinski, 1988 for detailed discussions and examples). For example, Schwarz, Strack, and Mai (in press) asked respondents to report their marital satisfaction as well as their satisfaction with life as a whole, under different order conditions. When the general life-satisfaction question preceded the marital satisfaction question, both measures were moderately correlated, $r = .32$. When the question order was reversed, however, the correlation increased to $r = .67$, $z = 2.61$, $p < .005$ for the difference between correlations (see also Strack, Martin, and Schwarz, 1988). This finding indicates that respondents were more likely to consider information bearing on their marriage in evaluating their life as a whole when this information was more accessible as a function of having answered the preceding question. Theoretically, one may assume that the likelihood of an early truncation of the search process increases with increasing time pressure (Kruglanski, 1980). Accordingly, carry-over effects as a function of the increased accessibility of previously used information should be more likely under telephone and face-to-face interview conditions than under self-administered conditions, due to the higher time pressure under the former administration modes.

In addition, the impact of a specific piece of highly accessible information has been shown to decrease as other, competing information becomes more accessible. The above-mentioned study on reports of life satisfaction (Schwarz *et al.*, in press) may again serve as an illustration. In one condition of this study, respondents were asked to report their satisfaction with three different life domains, including their marriage as the last question asked, prior to answering the general life-satisfaction question. Under this condition the increase in the correlation of marital satisfaction and general life-satisfaction was less pronounced, $r = .48$, than if respondents' marriage was the only life domain that was addressed, $r = .67$, $z = 1.57$, $p < .06$. This raises the possibility that accessibility effects resulting from preceding questions may be dampened under self-administered conditions because respondents may also read subsequent questions, thus increasing the accessibility of other, competing information. Accordingly, we may expect that carry-over effects as a function of the accessibility of information that was used in answering preceding questions may be more pronounced under telephone interview conditions than under self-administered questionnaire conditions. Unfortunately, controlled studies on the impact of administration mode on these processes are not yet available.

However, question order effects are not only a function of automatic accessibility processes, but may also emerge as a function of the deliberate consideration of the meaning of a question. In this regard it is important to note that respondents have more time and more opportunity to deliberately relate different questions—and their responses to them—to one another when the questions are presented in a self-administered questionnaire, rather than in a telephone or face-to-face interview. Accordingly, we may expect that responses to related questions are more consistent under self-administered conditions, and that respondents make more use of question context to determine the meaning of ambiguous questions. Accordingly, we may expect a self-administered format to increase the impact of question context—independently of the order in which the questions are asked—on the more or less 'conscious' process of question interpretation, but to decrease the impact of preceding

questions on the more or less 'automatic' process of carry-over effects. While data on the latter possibility are not yet available, the former possibility is well supported by several split-ballot experiments with German and U.S. samples, reported in detail by Bishop, Hippler, Schwarz, and Strack (1988) and Schwarz, Bishop, Hippler, and Strack (1989).

For example, Schuman and Ludwig (1983) assessed respondents' attitudes toward limiting Japanese imports to the U.S., and limiting U.S. exports to Japan, varying the order in which the questions were presented. Respondents were found to be more likely to favour limiting Japanese imports to the U.S. than they were to favour limiting U.S. exports to Japan when each question was asked in the first position. However, support for limiting U.S. exports to Japan increased when the question about it was asked *after* the question about restrictions on Japanese imports to the U.S., presumably because this question order evokes a norm of even-handedness (cf. Wyer and Srull, 1989, for a cognitive conceptualization of norms as implicational molecules).

Table 2 shows a conceptual replication of the Japanese trade items with a German sample (Bishop *et al.*, 1988), using telephone interview and self-administered questionnaire conditions. Respondents in the telephone interview condition were significantly more likely to favour limiting Japanese imports to Germany (36.5 per cent) than they were to favour limiting German exports to Japan (12.8 per cent), when each question was asked in the first position. Moreover, support for limitations on German exports to Japan increased (30.7 per cent) when this question was preceded by the question about restricting Japanese imports to Germany, which presumably evoked the norm of even-handedness. Thus, the data under telephone interview conditions replicate Schuman and Ludwig's (1983) findings based on a U.S. sample.

But when respondents were asked these same questions in the self-administered form, the *order* in which they were presented had—as expected—no significant effect on the results. Rather, these respondents reported considerable support for limiting German exports to Japan under *both* order conditions, suggesting that the norm of even-handedness was evoked independent of question order. This, of course, is what would be expected if respondents read both questions about trade restrictions before answering them. Data from a U.S. sample followed the same pattern, although the differences did not reach significance (see Bishop *et al.*, 1988).

In summary, these and related findings indicate that the impact of *question order* may be greatly reduced when the questions are presented in a self-administered questionnaire, which in effect eliminates sequential question presentation. The impact of *question context*, i.e. the impact of the *content* of adjacent questions, on the other hand, may be more pronounced under self-administered conditions and may emerge independently of question order, as is reflected in the operation of the norm of even-handedness in the data reviewed above. In combination, these data support the hypothesis that the simultaneous presentation of questions in a self-administered questionnaire may eliminate *order* effects but may enhance the impact of the *content* of related questions, because respondents have more time to think about their implications.

The latter assumption is further supported by another split-ballot experiment conducted as part of the same studies (see Schwarz *et al.*, 1989 for a more detailed report). Specifically, respondents were asked to report their attitudes toward a fictitious issue, namely the 'International Trade Act of 1986' (a question modelled after

Table 2. Question order effects as a function of mode of data collection (percentages)

	Telephone		Self-administered	
Germany/Japan	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	100.0	100.0	100.0
	(78)	(75)	(90)	(92)
	$\chi^2 = 7.34$, d.f. = 1, $p < .01$		$\chi^2 = 0.57$, d.f. = 1, n.s.	
Three-way interaction (response by form by mode)			$\chi^2 = 6.58$, d.f. = 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	100.0	100.0	100.0
	(78)	(74)	(90)	(92)
	$\chi^2 = 2.65$, d.f. = 1, $p = .103$		$\chi^2 = 1.07$, d.f. = 1, n.s.	
Three-way interaction (response by form by mode)			$\chi^2 = 3.64$, d.f. = 1, $p = .056$	

previous research by Schuman and Presser, 1981; Bishop, Oldendick, Tuchfarber, and Bennett, 1980; and Bishop, Oldendick, and Tuchfarber, 1986). As elaborated elsewhere (Schwarz and Strack, 1988; Strack and Martin, 1987), asking a question about an issue presupposes that the issue exists and places the burden on the respondent to determine its meaning. To do so respondents may either ask the interviewer for clarification or may depend on the context of the ongoing communication to disambiguate the question. While the former option is not available under self-administered questionnaire conditions, and does usually not result in helpful answers under standardized interview conditions either, self-administered questionnaire conditions do provide excellent opportunities to consult the context of the ambiguous question to make sense of it, whereas these opportunities are severely restricted under interview conditions. Accordingly, respondents under self-administered questionnaire conditions should be more likely to rely on the content of apparently related questions in answering an ambiguous one. In line with this assumption, responses to the International Trade Act question were found to be closely related to responses to the above import/export restriction items—asked five questions earlier—if the questions were presented in a self-administered questionnaire (gamma = .69), but not if they were presented in a telephone interview (gamma = .11).

In combination, these findings illustrate that it is important to distinguish between *question order effects* and *question context effects*, a distinction that has received little attention in previous research. While question order effects are likely to be reduced under self-administered questionnaire conditions, question context effects based on deliberate consideration of related questions are likely to be enhanced, due to respondents' increased opportunity to refer to preceding as well as subsequent

items. Obviously, these considerations are not limited to survey research but apply as well to the sequential or simultaneous presentation of items in laboratory experiments and psychological testing (see Schwarz and Sudman, in press, for examples).

Response order effects

The order in which response alternatives are presented to respondents has long been known to affect the obtained results (cf. Payne, 1951). Theoretically, *primacy effects*, that is, higher endorsements of items presented early in the list, as well as *recency effects*, that is, higher endorsements of items presented late in the list, may be obtained. While response order effects have occasionally been reported when the response alternatives present an ordered set of categories that constitute a verbal rating scale (e.g. excellent, very good, good, fair, poor), they are rare under these conditions (see Mingay and Greenwell, 1990). In contrast, response order effects have frequently been obtained when each response alternative presents a different opinion on an issue, and respondents are asked to select the one that best represents their own position. As a heuristic framework for understanding the nature of these latter effects, it has been suggested that each response alternative of that type can be portrayed as a single persuasive argument (Schwarz, Hippler, and Noelle-Neumann, in press). Borrowing from research on the processing of persuasive communications (see Petty and Cacioppo, 1986 for a detailed review), one may then assume that a given item is more likely to be endorsed the more positive cognitive responses it elicits, that is, the more agreeing thoughts the respondent generates. Conversely, a given item should be less likely to be endorsed the more disagreeing responses it elicits. The number of cognitive responses, however, is not only a function of the content of the item *per se*, but also a function of the degree of cognitive elaboration that a given mode of data collection permits.

Suppose, for example, that a long list of response alternatives is presented to respondents on a show-card as part of a face-to-face interview, or in a self-administered questionnaire. Under these conditions, 'items presented *early* in a list are likely to be subjected to deeper cognitive processing', as Krosnick and Alwin (1987, p. 213) noted. 'By the time a respondent considers the later alternatives, his or her mind is likely to be cluttered with thoughts about previous alternatives that inhibit extensive consideration of later ones.' Accordingly, a given response alternative is more likely to be endorsed if presented early rather than late in the list, provided that it is plausible to the respondent, thus eliciting agreeing thoughts. Conversely, an implausible response alternative, that elicits disagreeing thoughts, is less likely to be endorsed if presented early. Moreover, we may expect that these order effects are more pronounced under face-to-face interview conditions with show-cards than under self-administered questionnaire conditions, due to the differential time pressure under these administration modes. However, controlled experiments bearing on the latter assumption, and using complex response alternatives of the type described above, are not yet available.

Assume, however, that the items are not presented visually, but are read to respondents by the interviewer, either under face-to-face or telephone interview conditions. In this case respondents have little opportunity to elaborate on the items presented early in the list, because the time that is available for processing each item is restricted by the speed with which the interviewer moves on to read the next one. 'Under

these circumstances, respondents are able to devote most processing time to the *final* item(s) read, since interviewers usually pause most after reading them' (Krosnick and Alwin, 1987, p. 203). In addition, respondents may find it difficult to keep all response alternatives in mind without visual help. Accordingly, plausible items should be more likely to be endorsed if presented late rather than early in the list, resulting in recency effects under auditory presentation formats. Again, the reverse holds for implausible items.

In summary, response order effects are assumed to depend on the items' serial position, their plausibility, and the administration mode used. If the response alternatives are presented on show-cards or in a self-administered questionnaire, items presented early in the list are more likely to be extensively processed than items presented later, resulting in *primacy effects*, provided that the item is plausible to the respondents. In contrast, if the items are read to respondents, the last response alternatives are more likely to be extensively processed and recalled than the first ones, resulting in *recency effects*, again assuming plausibility of the items. Given that the likelihood of endorsement may be expected to decrease as more extensive processing uncovers flaws in implausible items, the reverse predictions hold for items that lack plausibility (see Schwarz *et al.*, in press, for a more detailed discussion). However, survey researchers typically avoid response alternatives that are likely to be implausible to a considerable number of respondents, thus limiting the currently available data to plausible response alternatives.

At present the best evidence for the predicted interaction of serial position and administration mode for plausible items comes from secondary analyses of a large number of split-ballot experiments with representative samples of the adult population in West Germany, that were originally conducted by the Allensbach Institute under the direction of Elizabeth Noelle-Neumann since the early 1950s. These experiments explored the impact of response order on answers to questions that involved any number from two to 30 response alternatives. Because long lists of response alternatives are typically *not* read to respondents, however, appropriate mode comparisons are not available for long lists (and are considered irrelevant by survey practitioners). Accordingly, we have to limit our analysis to dichotomous and trichotomous questions that were either presented on show-cards or read to respondents.

Overall, significant response order effects emerged on about 40 per cent (22 out of 54) of the questions that we could locate. When the response alternatives were presented on show-cards as part of face-to-face interviews, all order effects that reached significance were in the direction of *primacy effects*. For example, when asked whether government or private charity should play the dominant role in social welfare, 64 per cent of the respondents chose 'government' when this alternative was presented first, whereas only 50 per cent did so when this option was presented second. In contrast, *recency effects* were typically obtained when the response alternatives were read to respondents (see Schwarz *et al.*, in press). For example, when asked whether they would rather read a humorous or a serious novel, 48 per cent of the respondents chose the humorous novel when presented first, whereas 53 per cent did so when this alternative was presented second.

In addition, in some experiments a combination of visual and auditory presentation formats was used. Specifically, the response alternatives were read to respondents *before* they were presented on a show-card to facilitate the respondent's answer. In most experiments of this type, *recency effects* were likely to emerge, suggesting

that respondents process the response alternatives while they are read to them by the interviewer, without too much attention to either the accompanying or subsequent presentation of a show-card. Accordingly, the data pattern follows the pattern observed under a purely auditory administration mode.

Data bearing on response order effects in long list, on the other hand, are limited to a visual presentation format because researchers avoid reading excessively long lists to respondents. In most of these studies (see Krosnick and Alwin, 1987; Schwarz *et al.*, in press for details), primacy effects were observed, as would be expected on the basis of the present argument.

From an applied point of view, the most problematic implication of these findings is certainly that the direction of response order effects depends on the administration mode used. Most importantly, this finding indicates that mail surveys or face-to-face interviews with the help of show-cards may render results that are quite different from the results of telephone interviews without the use of show-cards, given that the primacy effects that emerge in one mode combine with the recency effects that emerge in the other. Note, however, that our conclusions are based on secondary analyses involving different questions under different administration modes. While the consistency of the data patterns across widely different questions suggests that the conclusions are likely to be valid, more tightly controlled experiments using the same questions and comparable samples under all conditions are definitely needed, as are experiments that explicitly manipulate item plausibility.

Retrospective reports: recall of information from memory

The recall of information from memory is known to improve with the amount of time that is available to search memory (cf. Anderson, 1980). Recalling specific events such as going out for a drink, for example, may take up to several seconds (Reiser, Black, and Abelson, 1985), and repeated attempts to recall may result in the retrieval of additional material, even after a considerable number of previous trials (e.g. Means and Loftus, in press and Williams and Hollan, 1981). Unfortunately, respondents are unlikely to have sufficient time to engage in repeated retrieval attempts in most survey situations (and may often not be motivated to do so even if they had the time).

Accordingly, recall should be poorest under telephone interview conditions, due to the high degree of time pressure under this mode of data collection, and best under self-administered questionnaire conditions, where respondents usually can take as much time as they like. For the same reason, differences due to respondents' motivation should be most pronounced under self-administered conditions, and should be least pronounced under telephone conditions, where the pressure of the situation is likely to override any desire to spend more time on the task.

In contrast, techniques that are designed to give the respondent more time to recall information from memory—e.g. increasing question length through the addition of redundant information (e.g. Blair, Sudman, Bradburn, and Stocking, 1977)—should prove irrelevant under self-administered questionnaire conditions but should affect the obtained responses under face-to-face and telephone conditions.

To test these hypotheses, Schwarz *et al.* (1989) asked German students to recall the year of the Falkland Islands war and the Soviet intervention in Afghanistan, under telephone interview as well as self-administered questionnaire conditions, and

in response to short or long, but redundant, questions. As expected, respondents who received self-administered questionnaire (under controlled classroom conditions that excluded the use of reference sources)[1] were significantly more likely to provide a correct answer to the Falkland Islands (43.3 per cent) and the Afghanistan question (46.2 per cent) than their classmates assigned to the telephone interview conditions (28.3 and 30.4 per cent, respectively).

Question length, on the other hand, did not affect the accuracy of recall under either administration mode. The unexpected absence of a question length effect under telephone interview conditions may either be due to the minimal time difference induced by the variation in question length or to the nature of the task. In the latter regard it is important to note that the *accuracy* of recall, rather than the *amount* of recalled material, was assessed, in contrast to previous studies (e.g. Blair *et al.*, 1977; Cannell, Miller, and Oksenberg, 1981). As previous research has shown, question length increases answer length, and it may well be that answer length, in turn, increases the likelihood that respondents generate helpful recall cues in the process of elaborating their answer (see Hippler and Schwarz, 1987 for a more detailed discussion). If only a short response, such as a specific date, is required, however, this potential advantage of longer questions may be eliminated—a possibility that deserves further research.

Social desirability and interviewer effects

The influence of social desirability depends significantly on the perceived anonymity of the responses (cf. DeMaio, 1984; Short, Williams, and Christie, 1976). In general, methods of data collection that provide high confidentiality of the response are likely to encourage fuller reports in response to threatening and sensitive behavioural questions. The available data by and large support this assumption (e.g. Colombotes, 1969; Hochstim, 1967; Knudsen, Pope, and Irish, 1967) but differences by method may disappear for extremely threatening questions, which are always substantially underreported (Bradburn and Sudman, 1979).

Regarding attitudinal responses, respondents have been found to be most likely to answer in a socially desirable fashion if the questions are asked in a face-to-face interview and least likely in a self-administered questionnaire (e.g. Smith, 1979; Strack, Schwarz, Chassein, Kern, and Wagner, in press). While telephone interviews may be expected to elicit intermediate social desirability concerns, due to the higher anonymity of the telephone interaction (see Frey, 1983, Ch. 2; Groves and Kahn, 1979; Short *et al.*, 1976), the available data are mixed and both more and less desirable responding has been observed under telephone conditions (see deLeeuw and van der Zouwen, 1988 for a review).

The direction of social desirability effects is likely to be mediated by interviewer

[1] Alternatively, one may assume that the classroom administration of the self-administered questionnaire reminded respondents of a test situation, resulting in an increased motivation to do well. We do not consider this possibility very compelling, because the questionnaire was introduced by an independent researcher and was not associated with either the content or the teacher of the respondents' class. Moreover, the recall questions were the only knowledge questions included in an opinion questionnaire, and were presented after a dozen opinion questions, under conditions of full anonymity. Finally, tests are not given as part of regular classes at German universities, but are restricted to specified time periods at the end of the semester. Thus, it seems highly unlikely that respondents misinterpreted the questionnaire administration as a testing situation.

characteristics and reactions (cf. Sudman and Bradburn, 1974). These are more likely to be noticed by respondents when they have face-to-face contact than when the interviewer cannot be seen, as is the case under telephone interview and self-administered questionnaire conditions. However, some interviewer characteristics that are known to affect socially desirable responding—such as age, sex (e.g. Moore, 1989) or race (e.g. Cotter, Cohen, and Coulter, 1982)—can be picked up from voice or speech characteristics under telephone interview conditions. Moreover, the interviewer's (dis)approval may be transmitted by paralinguistic variables that have been found to affect interviewer–respondent interaction (see Oksenberg and Cannell, 1988).

So far, we considered response effects that were expected to vary as a function of administration mode on the basis of theoretical considerations. We will now turn to response effects that may be less affected by the mode of data collection.

Wording and form effects in attitude questions

A large number of studies demonstrated that minor changes in the exact wording or form of a question can lead to major changes in the obtained responses (see Schuman and Kalton, 1985, and Schuman and Presser, 1981, for reviews). Three sources are likely to contribute to this phenomenon (cf. Hippler and Schwarz, 1987). First, changes in the wording of a question may result in changes in the question's substantive meaning. As Schuman and Kalton (1985) point out, most social issues are complex, 'yet individual survey questions must necessarily be kept simple' (p. 650). For this reason, different questions are likely to tap different facets of the same issue, resulting in different responses. Second, the same terms may mean different things to different people, as is illustrated in research by Belson (1968, 1981). Finally, changes in question wording or question form (such as the introduction of a middle alternative or a 'don't know' option) may affect what the respondent considers to be his or her task (e.g. Hippler and Schwarz, 1987, 1989). Because the information presented to respondents is essentially the same under all administration modes, each of these aspects can be assumed to be fairly independent of the mode of data collection used. The most notable exception to this assumption is provided by ambiguous or highly complicated question wordings, which may require extensive deliberation about the question's meaning, as we have discussed in the context of question order effects.

Question wording

The most pronounced, and probably the most frequently studied, question wording effect in the survey literature is the forbid-allow asymmetry, originally introduced by Rugg (1941). Respondents are either asked if something should be 'forbidden' (yes or no), or if it should be 'allowed' (yes or no). Given this format, respondents are more likely to say that it should 'not be forbidden' than that it should be 'allowed', and are more likely to say that it should 'not be allowed' than that it should be 'forbidden', even though both question forms appear to be logically equivalent.

An analysis of the cognitive processes underlying this phenomenon (Hippler and Schwarz, 1986) suggests that respondents focus on the implications of *doing* what they are asked about, namely forbidding or allowing something, rather than on the implications of *not* doing it. For this reason, indifferent respondents respond

'no' to *both* question forms because they neither want to take a position in favour ('allow') of the issue they feel indifferent about, nor a position in opposition ('forbid') to that issue. Respondents who hold a clear pro or con position, on the other hand, are not influenced by the wording of the question (see Bishop, 1989, for a replication of this finding).

As may be expected on the basis of these assumptions, a comparison of the forbid-allow asymmetry under self-administered questionnaire and telephone interview conditions indicated that neither its emergence nor its strength depended on the mode of data collection used (Bishop *et al.*, 1988). To which degree this finding generalizes to other wording effects remains an open issue.

Question form

Similarly, a number of question form effects was found to be independent of mode of data collection. In telephone interviews, as well as in self-administered questionnaires, respondents were more likely to select a *middle response alternative*, or a *no-opinion* alternative, if it was explicitly offered to them than if it had to be volunteered (Bishop *et al.*, 1988), replicating previous findings obtained under face-to-face interview conditions (e.g. Bishop, 1987; Schuman and Presser, 1981). Moreover, comparisons of *open* and *closed forms* of a question on work values also indicated no impact of administration mode. Under telephone interview as well as self-administered questionnaire conditions, most respondents who were given a closed question selected one of the precoded response alternatives and did not offer additional answers. Accordingly, the responses to the open question were considerably more heterogeneous than the responses to the closed question, again replicating previous findings (Schuman and Presser, 1981). However, responses to open-ended questions have been found to be shorter under telephone than under face-to-face interview conditions (e.g. Groves and Kahn, 1979; Sykes and Collins, 1988), which is usually attributed to the faster pace of telephone interviews and the absence of encouraging non-verbal feedback.

In combination, these and related findings (see Bishop *et al.*, 1988) suggest that most question wording and question form effects are likely to be relatively independent of the mode of data collection used.

CONCLUSIONS

In summary, the present review suggests that the mode of data collection is likely to affect variables that are known to mediate response effects in survey interviews (see Strack and Martin, 1987). The process of question answering first requires respondents to understand the meaning of the question. To interpret the question, respondents may refer to the content of apparently related questions. Under face-to-face and telephone interview conditions this possibility is restricted to preceding questions. Under self-administered conditions, on the other hand, respondents may also use the content of subsequent questions to disambiguate the meaning of preceding ones. Accordingly, context effects on the interpretation of ambiguous questions are order-dependent under the former administration modes but not under the latter (Bishop *et al.*, 1988). The interpretation of unambiguous questions, on the other hand, seems

to be relatively independent of the mode of data collection, and interaction effects of administration mode and question wording have so far not been demonstrated.

After having interpreted the question, respondents have to retrieve relevant information from memory to compute a judgement on the spot, unless a previously formed judgement can be recalled. In doing so they are unlikely to retrieve all potentially relevant information. Rather, they will truncate the search process as soon as enough information has come to mind to form a judgement (Bodenhausen and Wyer, 1987). How soon the retrieval process is truncated is a function of respondents' motivation and the time they have available to search memory (Kruglanski, 1980). Assuming sufficient motivation, respondents may be expected to engage in more extended retrieval efforts under self-administered than under face-to-face conditions. In contrast, they are likely to truncate the search process most quickly under the increased time pressure of telephone interview conditions. Accordingly, reliance on easily accessible information and the use of heuristic judgemental strategies (Wyer and Srull, 1989) may be expected to be most pronounced under the latter mode of data collection, which has also been found to result in less accurate recall of public events.

In addition to the information that respondents recall from memory, their judgements are a function of the thoughts that are elicited by the response alternatives presented to them, as conceptualized in the elaboration likelihood model of persuasion (Petty and Cacioppo, 1986). However, respondents' opportunity to elaborate on the response alternatives again depends on the administration mode. If the response alternatives are read to respondents, alternatives that are presented at the end of the list are more likely to be elaborated than alternatives presented earlier, whereas the reverse holds true if the response alternatives are presented in a visual format, either on show-cards as part of face-to-face interviews or in a self-administered questionnaire. As a result we find that plausible response alternatives, that elicit agreeing thoughts, are more likely to be endorsed if presented at the end of the list under an auditory presentation format, but at the beginning of the list under a visual presentation format (Krosnick and Alwin, 1987; Schwarz *et al.*, in press). In summary, the mode of data collection may influence respondents' judgemental processes via its impact on the retrieval of relevant information from memory, its impact on respondents' elaboration of the response alternatives presented to them, and its impact on the judgemental strategies used.

After having formed a judgement, respondents may need to format their judgement to fit the response alternatives provided by the researcher. The available data that bear on the impact of formal characteristics of the question (Bishop *et al.*, 1988) suggest that these processes are likely to be independent of the administration mode used, although responses to open-ended questions have been found to be shorter under telephone interview conditions (e.g. Groves and Kahn, 1979).

Finally, respondents need to communicate their judgement to the researcher. Not surprisingly, respondents have been found to be most likely to provide socially desirable responses under face-to-face interview conditions, and least likely to do so under self-administered conditions (e.g. Smith, 1979), whereas the data bearing on telephone interview conditions are mixed (deLeeuw and van der Zouwen, 1988).

Although many of the variables discussed in the introduction to the present paper have not yet been addressed in empirical research, the available evidence suggests that the impact of administration mode on some of the better-documented response effects in survey measurement can be plausibly conceptualized on the basis of psycho-

logical considerations. This provides encouraging evidence for the usefulness of a psychological approach to survey methodology that incorporates cognitive and communicational variables (Hippler *et al.*, 1987; Jabine *et al.*, 1984; Schwarz and Sudman, in press). More importantly, it encourages a fuller exploitation of the theoretical literature in cognitive and social psychology than could be provided in the present paper. In particular, a more detailed application of theorizing and research in reading, listening, text comprehension, and discourse processing is likely to contribute to a deeper understanding of the issues raised in the present paper, and awaits further research. Without the systematic development and testing of guiding theoretical principles, our knowledge about survey measurement is 'likely to remain a set of scattered findings, with repeated failure at replication of results', as Groves and Lyberg (1988, p. 210) recently noted in a related discussion. We hope that the present preliminary outline of some of the key issues will contribute to the development of a more coherent body of knowledge about the psychological processes that underlie response effects in survey measurement and their dependency on the mode of data collection.

ACKNOWLEDGEMENTS

The reported research was supported by grant SWF0044-6 from the Bundesministerium für Forschung und Technologie to Norbert Schwarz, grant Str 264/2 from the Deutsche Forschungsgemeinschaft to Fritz Strack and Norbert Schwarz, and ZUMA's programme on cognition and survey research. We want to thank Tom Ostrom and an anonymous reviewer for their comments on a previous draft of this paper. Correspondence should be addressed to Norbert Schwarz, ZUMA, PO Box 12 21 55, D-6800 Mannheim, West Germany.

REFERENCES

Anderson, J. R. (1980). *Cognitive psychology and its implications*. San Francisco: Freeman.

Argyle, M. (1969). *Social interaction*. London: Methuen.

Ball, D. W. (1968). Toward a sociology of telephones and telephoners. In M. Truzzi (ed.), *Sociology and everyday life* (pp. 59–75). Englewood Cliffs, NJ: Prentice-Hall.

Belson, W. A. (1968). Respondent understanding of survey questions. *Polls*, 3, 1–13.

Belson, W. A. (1981). *The design and understanding of survey questions*. Aldershot: Gower

Biemer, P. P. (1988). Measuring data quality. In R. M. Groves, P. Biemer, L. Lyberg, J. Massey, W. Nicholls, and J. Waksberg (eds), *Telephone survey methodology* (pp. 273–282). New York: Wiley.

Bishop, G. F. (1989). Another look at the forbid-allow asymmetry. Paper presented at the Annual Conference of the American Association for Public Opinion Research, St. Petersburg, Florida (May).

Bishop, G. F., Hippler, H. J., Schwarz, N. and Strack, F. (1988). A comparison of response effects in self-administered and telephone surveys. In R. M. Groves, P. Biemer, L. Lyberg, J. Massey, W. Nicholls, and J. Waksberg (eds), *Telephone survey methodology* (pp. 321–340). New York: Wiley.

Bishop, G. F., Oldendick, R. W. and Tuchfarber, A. J. (1986). Opinions on fictitious issues: the pressure to answer survey questions. *Public Opinion Quarterly*, **50**, 240–250.

Bishop, G. F., Oldendick, R. W., Tuchfarber, A. J. and Bennett, S. E. (1980). Pseudo-opinions on public affairs. *Public Opinion Quarterly*, **44**, 198–209.

Blair, E., Sudman, S., Bradburn, N. M. and Stocking, C. B. (1977). How to ask questions

about drinking and sex: response effects in measuring consumer behavior. *Journal of Marketing Research*, **14**, 316–321.

Bodenhausen, G. V. and Wyer, R. S. (1987). Social cognition and social reality: information acquisition and use in the laboratory and the real world. In H. J. Hippler, N. Schwarz, and S. Sudman (eds), *Social information processing and survey methodology* (pp. 6–41). New York: Springer-Verlag.

Bradburn, N. M., Rips, L. J. and Shevell, S. K. (1987). Answering autobiographical questions: the impact of memory and inference on surveys. *Science*, **236**, 157–161.

Bradburn, N., Sudman, S. and Associates (1979). *Improving interviewing methods and questionnaire design*. San Francisco: Jossey-Bass.

Brown, N. R., Rips, L. J. and Shevell, S. K. (1985). The subjective dates of natural events in very-long-term memory. *Cognitive Psychology*, **17**, 139–177.

Brown, N. R., Rips, L. J. and Shevell, S. K. (1986). Public memories and their personal context. In D. C. Rubin (ed.), *Autobiographical memory* (pp. 137–158). Cambridge: Cambridge University Press.

Cannell, C. F., Miller, P. V. and Oksenberg, L. (1981). Research on interviewing techniques. In S. Leinhardt (ed.), *Sociological Methodology 1981*. San Francisco: Jossey-Bass.

Cotter, P. R., Cohen, J. and Coulter, P. B. (1982). Race-of-interviewer effects in telephone interviews. *Public Opinion Quarterly*, **46**, 278–284.

DeLeeuw, E. D. and van der Zouwen, J. (1988). Data quality in telephone and face-to-face surveys: a comparative analysis. In R. M. Groves, P. Biemer, L. Lyberg, J. Massey, W. Nicholls, and J. Waksberg (eds), *Telephone survey methodology* (pp. 283–293). New York: Wiley.

DeMaio, T. J. (1984). Social desirability and survey measurement: a review. In C. F. Turner and E. Martin (eds), *Surveying subjective phenomena*, Vol. 2 (pp. 257–281). New York: Russell-Sage.

Dijkstra, W. and van der Zouwen, J. (eds) (1982). *Response behavior in the survey interview*. New York: Academic Press.

Dillman, D. (1978). *Mail and telephone surveys: the total design method*. New York: Wiley.

Dillman, D. A. and Tarnai, J. (1988). Administrative issues in mixed mode surveys. In R. M. Groves, P. Biemer, L. Lyberg, J. Massey, W. Nicholls, and J. Waksberg (eds), *Telephone survey methodology* (pp. 509–528). New York: Wiley.

Frey, J. H. (1983). *Survey research by telephone*. Newbury Park, CA: Sage.

Groves, R. M. and Kahn, R. L. (1979). *Surveys by telephone: A national comparison with personal interviews*. New York: Academic Press.

Groves, R. M. and Lyberg, L. E. (1988). An overview of nonresponse issues in telephone surveys. In R. M. Groves, P. Biemer, L. Lyberg, J. Massey, W. Nicholls, and J. Waksberg (eds), *Telephone survey methodology* (pp. 191–211). New York: Wiley.

Hippler, H. J. and Schwarz, N. (1986). Not forbidding isn't allowing: the cognitive basis of the forbid-allow asymmetry. *Public Opinion Quarterly*, **50**, 87–96.

Hippler, H. J. and Schwarz, N. (1987). Response effects in surveys. In H. J. Hippler, N. Schwarz, and S. Sudman (eds), *Social information processing and survey methodology* (pp. 102–122). New York: Springer-Verlag.

Hippler, H. J. and Schwarz, N. (1989). 'No opinion' filters: a cognitive perspective. *International Journal of Public Opinion Research*, **1**, 77–87.

Hippler, H. J., Schwarz, N. and Sudman, S. (eds) (1987). *Social information processing and survey methodology*. New York: Springer-Verlag.

Hochstim, J. R. (1967). A critical comparison of three strategies of collecting data from households. *Journal of the American Statistical Association*, **62**, 976–989.

Jabine, T. B., Straf, M. L., Tanur, J. M. and Tourangeau, R. (eds) (1984). *Cognitive aspects of survey methodology: building a bridge between disciplines*. Washington, DC: National Academy Press.

Kahneman, D., Slovic, P. and Tversky, A. (1982). *Judgment under uncertainty: heuristics and biases*. Cambridge: Cambridge University Press.

Knudsen, D. D., Pope, H. and Irish, D. P. (1967). Response differences to questions on sexual standards: an interview-questionnaire comparison. *Public Opinion Quarterly*, **31**, 290–297.

Krosnick, J. A. and Alwin, D. F. (1987). An evaluation of a cognitive theory of response order effects in survey measurement. *Public Opinion Quarterly*, **51**, 201–219.

Kruglanski, A. W. (1980). Lay epistemologic process and contents: another look at attribution theory. *Psychological Review*, **87**, 70–87.

Means, B., Mingay, D. J., Nigam, A. and Zarrow, M. (1988). A cognitive approach to enhancing health survey reports of medical visits. In M. M. Gruneberg, P. E. Morris, and R. N. Sykes (Eds) *Practical aspects of memory: current research and issues*, Vol. 1 (pp. 537–542). Chichester: Wiley.

Mingay, D. J. and Greenwell, M. T. (1990). Memory bias and response order effects. *Journal of Official Statistics*.

Moore, D. (1989). Effects of interviewers' sex on reported rates of child abuse and attitudes toward child-rearing practices. Paper presented at the Annual Conference of the American Association for Public Opinion Research, St. Petersburg, Florida (May).

Payne, S. L. (1951). *The art of asking questions*. Princeton, NJ: Princeton University Press.

Petty, R. E. and Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change*. New York: Springer-Verlag.

Reid, A. (1977). Comparing telephone with face-to-face contact. In I. Pool (ed.), *The social impact of the telephone*. Cambridge, MA: MIT Press.

Reiser, B. J., Black, J. B. and Abelson, R. P. (1985). Knowledge structure in the organization and retrieval of autobiographical memories. *Cognitive Psychology*, **17**, 89–137.

Rogers, T. F. (1976). Interviews by telephone and in person: quality of responses and field performance. *Public Opinion Quarterly*, **40**, 51–65.

Rugg, D. (1941). Experiments in wording questions: II. *Public Opinion Quarterly*, **5**, 91–92.

Schuman, H. and Kalton, G. (1985). Survey methods. In G. Lindzey and E. Aronson (eds), *Handbook of social psychology*, Vol. 1 (pp. 635–697). New York: Random House.

Schuman, H. and Ludwig, J. (1983). The norm of even-handedness in surveys as in life. *American Sociological Review*, **48**, 112–120.

Schuman, H. and Presser, S. (1981). *Questions and answers in attitude surveys*. New York: Academic Press.

Schwarz, N. (1990). Assessing frequency reports of mundane behaviors: contributions of cognitive psychology to questionnaire construction. In C. Hendrick and M. S. Clark (eds), *Research methods in personality and social psychology* (Review of Personality and Social Psychology, Vol. 11, pp. 98–119). Beverly Hills, CA: Sage.

Schwarz, N., Bishop, G., Hippler, H. J. and Strack, F. (1989). A comparison of telephone and self-administered surveys: question context effects and the veridicality of retrospective reports. Under review.

Schwarz, N., Hippler, H. J. and Noelle-Neumann, E. (in press). A cognitive model of response order effects in survey measurement. In N. Schwarz and S. Sudman (eds), *Context effects in social and psychological research*. New York: Springer-Verlag.

Schwarz, N. and Strack, F. (1988). The survey interview and the logic of conversation. *ZUMA-Arbeitsbericht*, No. 3/1988.

Schwarz, N., Strack, F. and Mai, H. P. (in press). Assimilation and contrast effects in part-whole question sequences: A conversational logic analysis. *Public Opinion Quarterly*.

Schwarz, N. and Sudman, S. (eds) (in press). *Context effects in social and psychological research*. New York: Springer-Verlag.

Short, J., Williams, E. and Christie, B. (1976). *The social psychology of telecommunications*. London: Wiley.

Smith, T. W. (1979). Happiness. *Social Psychology Quarterly*, **42**, 18–30.

Strack, F., Erber, R. and Wicklund, R. (1982). Effects of salience and time pressure on ratings of social causality. *Journal of Experimental Social Psychology*, **18**, 581–594.

Strack, F. and Martin, L. (1987). Thinking, judging, and communicating: a process account of context effects in attitude surveys. In H. J. Hippler, N. Schwarz and S. Sudman (eds), *Social information processing and survey methodology* (pp. 123–148). New York: Springer-Verlag.

Strack, F., Martin, L. L. and Schwarz, N. (1988). Priming and communication: the social determinants of information use in judgments of life-satisfaction. *European Journal of Social Psychology*, **18**, 429–442.

Strack, F., Schwarz, N., Chassein, B., Kern, D. and Wagner, D. (in press). The salience of comparison standards and the activation of social norms: consequences for judgments of happiness and their communication. *British Journal of Social Psychology*.

Strube, G. (1987). Answering survey questions: the role of memory. In H. J. Hippler, N. Schwarz, and S. Sudman (eds), *Social information processing and survey methodology* (pp. 86–101). New York: Springer-Verlag.

Sudman, S. and Bradburn, N. M. (1974). *Response effects in surveys: a review and synthesis*. Chicago, IL: Aldine.

Sudman, S. and Bradburn, N. M. (1983). *Asking questions*. San Francisco, CA: Jossey-Bass.

Sykes, W. and Collins, M. (1988). Effects of mode of interview: Experiments in the UK. In R. M. Groves, P. Biemer, L. Lyberg, J. Massey, W. Nicholls, and J. Waksberg (eds), *Telephone survey methodology* (pp. 301–320). New York: Wiley.

Tourangeau, R. (1984). Cognitive sciences and survey methods. In T. B. Jabine *et al.* (eds), *Cognitive aspects of survey methodology: building a bridge between disciplines* (pp. 73–100). Washington, DC: National Academy Press.

Tourangeau, R. (1987). Attitude measurement: a cognitive perspective. In H. J. Hippler, N. Schwarz, and S. Sudman (eds), *Social information processing and survey methodology* (pp. 149–162). New York: Springer-Verlag.

Tourangeau, R. and Rasinski, K. A. (1988). Cognitive processes underlying context effects in attitude measurement. *Psychological Bulletin*, **103**, 299–314.

Williams, E. (1977). Experimental comparisons of face-to-face and mediated communication: a review. *Psychological Bulletin*, **84**, 963–976.

Williams, M. D. and Hollan, J. D. (1981). The process of retrieval from very long term memory. *Cognitive Science*, **5**, 87–119.

Wyer, R. S. and Srull, T. K. (1989). *Memory and cognition in its social context*. Hillsdale, NJ: Erlbaum.