

Assessing frequency reports of mundane behaviors: contributions of cognitive psychology to questionnaire construction

Schwarz, Norbert

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. (1988). *Assessing frequency reports of mundane behaviors: contributions of cognitive psychology to questionnaire construction*. (ZUMA-Arbeitsbericht, 1988/10). Mannheim: Zentrum für Umfragen, Methoden und Analysen -ZUMA-. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-70459>

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.

The attached reprint replaces ZUMA-Arbeitsbericht No. 88/10 by the same author.

Schwarz, Norbert. Assessing frequency reports of mundane behaviors: Contributions of cognitive psychology to questionnaire construction. In: C. Hendricks & M. Clark (Eds.), Research methods in personality and social psychology (Review of Personality and Social Psychology, Vol. 11). Beverly Hills: Sage, 1990.

REVIEW OF PERSONALITY AND SOCIAL PSYCHOLOGY

Editors

Clyde Hendrick, *Texas Tech University*

Margaret S. Clark, *Carnegie Mellon University*

Associate Editors

John H. Harvey, *University of Iowa*

Russell A. Jones, *University of North Florida*

Richard P. McGlynn, *Texas Tech University*

Walter G. Stephan, *New Mexico State University*

Editorial Board

Ellen Berscheid, *University of Minnesota*

Marilynn B. Brewer, *University of
California, Los Angeles*

Bram Buunk, *University of Nijmegen*

Donn Byrne, *SUNY, Albany*

Shelly Chaiken, *New York University*

Francis C. Danc, *Clemson University*

Keith E. Davis, *University of South
Carolina*

Valerian J. Derlega, *Old Dominion
University*

Steve Duck, *University of Iowa*

Alice H. Eagly, *Purdue University*

Alan C. Elms, *University of California,
Davis*

Kenneth J. Gergen, *Swarthmore College*

Mary Gergen, *Penn State/Delaware Cty*

J. Martin Giesen, *Mississippi State
University*

Lewis Goldberg, *Oregon Research
Institute*

Leo Goldberger, *New York University*

Judith A. Hall, *Northeastern University*

Susan S. Hendrick, *Texas Tech University*

Ted Huston, *University of Texas, Austin*

Warren H. Jones, *University of Tulsa*

Norbert L. Kerr, *Michigan State
University*

Ellen Langer, *Harvard University*

John M. Levine, *University of Pittsburgh*

J. W. Mann, *University of Witwatersrand*

Leon Mann, *Flinders University*

Martha T. S. Mednick, *Howard
University*

Lynn Carol Miller, *Scripps College*

Louls A. Penner, *University of South
Florida*

Daniel Perlman, *University of British
Columbia*

Marylyn Rands, *Wheaton College*

Harry T. Reis, *University of Rochester*

Deborah Richardson, *University of
Georgia*

Brendan Gail Rule, *University of Alberta*

Caryl E. Rusbult, *University of North
Carolina*

Judith C. Schwartz, *University of
California, Santa Cruz*

Phillip Shaver, *SUNY, Buffalo*

Cecilia H. Solano, *Wake Forest
University*

Cookie White Stephan, *New Mexico
State University*

Gifford Weary, *Ohio State University*

Daniel M. Wegner, *Trinity University*

A. Rodney Wellens, *University of Miami*

Ladd Wheeler, *University of Rochester*

Jerry S. Wiggins, *University of British
Columbia*

Mark P. Zanna, *University of Waterloo*

Research Methods in Personality and Social Psychology

Editors

CLYDE HENDRICK
MARGARET S. CLARK

11

REVIEW of PERSONALITY and SOCIAL PSYCHOLOGY

Published in cooperation with the Society for Personality and Social Psychology, Inc.



SAGE PUBLICATIONS

The Publishers of Professional Social Science

Newbury Park London New Delhi

1990

Assessing Frequency Reports of Mundane Behaviors

CONTRIBUTIONS OF COGNITIVE PSYCHOLOGY TO QUESTIONNAIRE CONSTRUCTION

NORBERT SCHWARZ

Norbert Schwarz received degrees in sociology (Dipl.-Soz., Dr. phil.) from the University of Mannheim and in psychology (Dr. phil. habil.) from the University of Heidelberg. He has held visiting appointments at North American universities, and he currently directs a research program on cognitive aspects of survey methodology at the Zentrum für Umfragen, Methoden und Analysen (ZUMA), in Mannheim, Federal Republic of Germany, and is Privatdozent of Psychology at the University of Heidelberg. His research interests focus on human judgmental processes, in particular the interplay of affect and cognition, and the application of social cognition research to survey methodology.

Much of our knowledge about individuals' behavior is based on their direct verbal reports. From consumer behavior to health problems, and from styles of parenting to the nation's unemployment rate or the prevalence of crime, psychologists and social scientists rely on respondents' behavioral reports as their major data base for testing theories of human behavior and offering advice on public policy. Given the importance of behavioral reports, surprisingly little attention has been given to how respondents go about answering quantitative autobiographical questions, asking them, for example, how often they have done something, or how much of something they have consumed, during a specified time period. As an introduction to the issues raised in the present chapter, the reader may want to answer the following questions, which are taken from major U.S. surveys:

1. During the two-week [reference] period, on the days when you drank liquor, about how many drinks did you have? (Health Interview Survey Supplement, National Center for Health Statistics)

AUTHOR'S NOTE: Parts of the research reported here were supported by grants Schw 278/2 and Str 264/3 from the Deutsche Forschungsgemeinschaft to the author and Fritz Strack, by a Feodor Lynen Fellowship from the Alexander von Humboldt Foundation to the author, and by ZUMA's program on cognition and survey research. Thanks are extended to Barbara Bickart, Bettina Scheuring, and Fritz Strack for their helpful comments on a previous draft.

2. Now, I'd like to read you a short list of different kinds of pain. Please say for each one, on roughly how many days — if any — in the last 12 months you have had that type of pain. How many days in the last year have you had [headaches; backaches; stomach pains; joint pains; muscle pains; dental pains]? (Health Interview Survey Supplement, National Center for Health Statistics)
3. We would like to ask [your/your relative's] work history for every job and every period of unemployment that lasted 6 months or more, starting with the first full-time job [you/he] held from the age of 16, including any jobs held while in military service. [. . .] About how many hours a week did (you/he) work on that [first] job, including overtime? ("Women at Work" Study, U.S. Bureau of the Census)
4. When you were growing up, how frequently did your father attend religious services? (General Social Survey, National Opinion Research Center)

The present chapter reviews our current knowledge of the cognitive processes that underlie respondents' answers to quantitative questions such as the ones above, paying special attention to the methodological implications of the reviewed findings for questionnaire construction. Much of the research is fairly new, growing out of the recently initiated dialogue between survey researchers and cognitive psychologists (see Hippler, Schwarz, & Sudman, 1987; Jabine, Straf, Tanur, & Tourangeau, 1984, for reviews), as well as the recent interest in autobiographical memory (see Gruneberg, Morris, & Sykes, 1988; Neisser, 1982; Rubin 1986).

THE PROCESS OF QUESTION ANSWERING

Answering a quantitative autobiographical question requires that respondents undertake several tasks. First, respondents need to understand what the question refers to, and which behavior they are supposed to report. Second, they have to recall or reconstruct relevant instances of this behavior from memory. Third, if the question specifies a reference period, they must determine if these instances occurred during this reference period or not. Similarly, if the question refers to their "usual" behavior, respondents have to determine if the recalled or reconstructed instances are reasonably representative or if they reflect a deviation from their usual behavior. Fourth, as an alternative to recalling or reconstructing instances of the behavior, respondents may rely on their general knowledge, or other salient information that may bear on their task, to infer an answer. Finally, respondents have to provide their report to the researcher. They may need to map their report onto a response scale provided to them, and they may want to edit it for rea-

sons of social desirability (see Strack & Martin, 1987; Tourangeau & Rasinski, 1988, for related discussions of attitude questions).

In the first part of this chapter, each of these tasks will be discussed in some detail. For the sake of simplicity, this discussion will assume that the question is asked in an open response format, and that no precoded response alternatives are provided to the respondent. As the review develops, it will become evident that answering autobiographical questions is a process that is highly theory-driven, and relies as much on respondents' inference strategies as on their recall of specific autobiographical details. The second part of the chapter will then explore a specific set of inference rules, namely, respondents' assumptions about the informational value of response alternatives provided to them by the researcher. Focusing on what respondents learn from response alternatives, this section will address the special issues that arise at various stages of the judgment process when respondents are asked to report their behavior by checking the appropriate alternative from a set of response categories in a closed answer format.

ANSWERING BEHAVIORAL QUESTIONS IN AN OPEN RESPONSE FORMAT

Understanding the Question

The key issue at the question-comprehension stage is whether the behavior that the respondent identifies as the referent of the question does or does not match what the researcher had in mind. As a general rule, question comprehension is considerably poorer than most researchers would like to believe, even for apparently simple questions (Belson, 1981). For example, in a British readership survey (Belson, 1968), respondents were presented an aided recall task: "I want you to go through this booklet with me, and tell me, for each paper, whether you happen to have looked at any copy of it in the past three months, it doesn't matter where" (p. 2). Subsequently, respondents were asked to explain the key terms of the question. The interpretations of "looked at" ranged from "seen on a bookstall" to "read fairly fully," and included "like to read," "bought," and "taken regularly." The phrase "any copy" was frequently overlooked or interpreted to mean "your own copy," and the reference period of three months tended to be ignored or became a vague "recently." Belson concluded that respondents who find a question difficult to answer are "likely to modify it in such a way that it becomes more easy to answer" (p. 9). In particular, they are likely to interpret broad terms

less broadly than intended, and to respond to the gist of the question rather than to its exact wording. Moreover, respondents may use the response alternatives provided to them by the researcher to determine the meaning of the question, as will be elaborated below.

In addition to these general comprehension problems (see Belson, 1981; Hippler & Schwarz, 1987; Strack & Martin, 1987 for reviews of related research), respondents' definition of certain behaviors may vary depending upon *who* is supposed to engage in the behavior. For example, the concept of "paid work," which is crucial for most labor force surveys, was found to include very different activities depending on the target person's age, education, and employment history. What qualifies as "paid work" for the respondent's teenage children (e.g., baby-sitting or mowing the neighbor's lawn) does not qualify as "paid work" for the adult respondent, resulting in different definitions of the *class* of target behaviors as a function of the employment history of the target person (Schwarz, 1987). Thus the same question may result in the assessment of different behaviors, rendering the reports noncomparable.

Recalling Relevant Instances

Once respondents have formed a subjective understanding of what the question refers to, they need to retrieve relevant information about the behavior under study from memory. Ideally, most researchers would like the respondent to scan the reference period, retrieve all instances that match the target behavior, and count them in order to determine the overall frequency of the behavior during the reference period. This, however, is the route that respondents are least likely to take.

In fact, except for rare and very important behaviors, respondents are unlikely to have detailed representations of numerous individual instances of a behavior stored in memory, and may be expected to blend details of various instances into one global representation of the behavior under study (Linton, 1982; Means, Mingay, Nigam, & Zarrow, 1988; Neisser, 1986; Wickelgren, 1976). Thus many individual episodes become indistinguishable or irretrievable, due to interference from other similar instances (Baddeley & Hitch, 1977; Wagenaar, 1986), fostering the generation of knowledgelike representations that "lack specific time or location indicators" (Strube, 1987, p. 89). The finding that a single spell of unemployment is more accurately recalled than multiple spells (Mathiowetz, 1986), for example, suggests that this phenomenon applies not only to mundane and unimportant behaviors, but also to repeated experiences that profoundly affect an individual's life.

Accordingly, a "recall and count" model does not provide an appropriate description of how people answer frequency questions about frequent behaviors or experiences. Rather, their answers are likely to be based on some fragmented recall and the application of inference rules to compute a frequency estimate, as will be described below (see Blair & Burton, 1987; Bradburn, Rips, & Shevell, 1987; Means et al., 1988).

If researchers are interested in obtaining reports that are based on recalled episodes, they may simplify respondents' task by providing appropriate recall cues. This, however, is more easily said than done. Theoretically, the most efficient recall cues are cues that match respondents' encodings of the target behavior. But, unfortunately, research on everyday memory is a relatively recent field of study and little is known about respondents' habits of encoding for specific classes of everyday events. For example, while researchers may be interested in how often respondents drink alcoholic beverages, respondents may be unlikely to encode "drinking alcoholic beverages" as a separate category. Therefore, providing them with a selection of common situations (seeing friends, watching TV, having dinner, and so on) in which they may consume alcohol may provide better recall cues (Strube, 1987). In the absence of specific knowledge about the representation of everyday behaviors, breaking down a global question into several more specific ones and using short rather than long reference periods were found to improve respondents' recall and to increase the likelihood that respondents used a recall rather than an estimation strategy.

For example, a question about how often a respondent has eaten out may be broken down into a series of separate questions about eating at different types of restaurants. A recent experiment indicated that when respondents were asked to report separately how often they had eaten dinner in Chinese, Greek, Italian, American, Mexican, and fast-food restaurants, they reported an average of 26.0 trips in a three-month period, compared to 20.5 trips reported when they were asked the more general question, "How many times have you eaten dinner in a regular or fast-food restaurant?" (Sudman & Schwarz, 1987). Moreover, when respondents were asked to report restaurant visits only for the previous month, they reported about 13.7 visits, a monthly average substantially greater than the range of 7-9 visits found by dividing the three-month totals by three.

In addition, varying the reference period and the specificity of the question affected respondents' strategies, as was previously reported by Blair and Burton (1987). When asked only about the previous month, most respondents (62%) reported that they tried to remember specific events. For the

three-month period, only about one-third (32%) of those asked the general question about all restaurants tried to remember specific events, while slightly more than half (56%) reported using an estimation strategy. For respondents asked about specific types of restaurants for three months, the majority (54%) reported using a mixed strategy. Specifically, they tried to remember specific trips to some kinds of restaurants—namely, the ones rarely visited—and estimated for other types. Thus respondents were more likely to use an estimation strategy the larger the number of similar experiences, presumably because repeated similar experiences are difficult to retrieve individually.

In general, the quality of recall will improve as the retrieval cues presented in the body of the question, or on an accompanying list of examples, become more specific. There is, however, an important drawback to the use of specific questions: Respondents are likely to omit instances that do not match the specific questions or examples, resulting in underreports if the list is not exhaustive. Thus, in the above example, visits to a restaurant not included in the set of specific questions are likely to be omitted even if a final question asks if the respondent ate at any "other" restaurant (see Sudman & Bradburn, 1983).

In providing specific recall cues, it is important to note that different cues are differentially effective. Thus the date of an event is usually found to be a poor cue, whereas cues pertaining to what happened, where it happened, and who was involved have been found to be very effective (Wagenaar, 1986, 1988). In addition, recall will improve when respondents are given sufficient time to search memory. Recalling specific events, such as going out for a drink, may take up to several seconds (Reiser, Black, & 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 et al., 1988; Williams & Hollan, 1981). Unfortunately, respondents are unlikely to have sufficient time to engage in repeated retrieval attempts in most research situations (and may often not be motivated to do so even if they had the time). This is particularly crucial in the context of survey research, where the available time per question is usually less than one minute (Bradburn et al., 1987; Groves & Kahn, 1979).

Moreover, the direction in which respondents search memory has been found to influence the quality of recall. Specifically, better recall is achieved when respondents begin with the most recent occurrence of a behavior and work backward in time than when they begin at the beginning of the reference period (e.g., Loftus & Fathi, 1985; Whitten & Leonard, 1981). This

presumably occurs because memory for recent occurrences is richer and the recalled instances may serve as cues for recalling previous ones. Given free choice, however, respondents tend to prefer the less efficient strategy of forward recall. Even under optimal conditions, however, respondents will frequently be unable to recall an event or some of its critical details, even if they believed they would "certainly" remember it at the time it occurred (e.g., Linton, 1975; Thompson, 1982; Wagenaar, 1986). In general, the available evidence suggests that respondents are likely to underreport behaviors and events, which has led many researchers to assume that higher reports of mundane behaviors are likely to be more valid. Accordingly, a "the more the better" rule is frequently substituted for external validity checks.

Dating Recalled Instances

After recalling or reconstructing a specific instance of the behavior under study, respondents have to determine if this instance occurred during the reference period. This requires that they understand the extension of the reference period and that they can accurately date the instance with regard to that period.

Reference periods that are defined in terms of several weeks or months are highly susceptible to misinterpretations. For example, the phrase "during the last twelve months" has been found to be construed as a reference to the last calendar year, as including or excluding the current month, and so on. Similarly, anchoring the reference period with a specific date—for example, "Since March 1, how often . . . ?"—is not very helpful because respondents will usually not be able to relate an abstract date to meaningful memories.

Not surprisingly, the most efficient way to anchor a reference period is the use of salient personal or public events, often referred to as "landmarks" (Loftus & Marburger, 1983). In addition to improving respondents' understanding of the reference period, the use of landmarks facilitates the dating of recalled instances. Given that the calendar date of an event will usually not be among its encoded features, respondents were found to relate recalled events to other, more outstanding events in order to reconstruct the exact time and day (e.g., Baddeley, Lewis, & Nimmo-Smith, 1978). Accordingly, using public events (such as the eruption of Mount Saint Helens), important personal memories that respondents were asked to think of, or outstanding dates (such as New Year's Eve) as landmarks was found to reduce dating biases (Loftus & Marburger, 1983; Means et al., 1988). A related procedure, called "bounded recall" (Neter & Waksberg, 1964), has been developed for use in repeated interviews, where data from the previous interview

serve as landmarks and recall cues. However, the high cost of repeated interviews frequently discourages its use, although Sudman, Finn, and Lannom (1984) adapted the procedure for use in single interviews.

Without a chance to relate a recalled event to a well-dated landmark or a personal "time line" (Means et al., 1988), time dating is likely to reflect both "forward" and "backward telescoping." That is, distant events are assumed to have happened more recently than they did, whereas recent events are assumed to be more distant than they are (e.g., Brown, Rips, & Shevell, 1985, 1986). Moreover, respondents have been found to use the clarity and vividness of their memory as a cue to the distance of the event, assuming that events that are recalled in detail occurred more recently than events for which memory is more impoverished (Brown, Rips, & Shevell, 1985, 1986). This heuristic fosters forward telescoping for sensational and vivid events.

The Role of Inference Processes

Given the inappropriateness of the "recall and count" model, it is not surprising that inference strategies play a major role in answering frequency questions. As Bradburn et al. (1987) observed, "Respondents will use any information they have in order to generate a reasonable answer" (p. 160). The best documented strategies are the use information provided by precoded response alternatives (discussed in detail in the second part of this chapter), the decomposition of the recall problem into subparts, the use of the availability heuristic, and reliance on subjective theories of stability and change.

Regarding *decomposition strategies*, respondents who were asked to report the number of restaurant visits were found to determine a rate of occurrence for a limited time period and to multiply this rate to arrive at an estimate for the complete reference period (Blair & Burton, 1987; Sudman & Schwarz, 1987). For example, a hypothetical respondent may first determine that she eats out about every weekend, and that she also had dinner at a restaurant this Wednesday, but apparently not the week before. Thus she may infer that this makes four times a month for the weekends, and let's say twice for other occasions, thus "eighteen times during the last three months" would be an appropriate answer. Note that estimates of this type are likely to be accurate if the respondent's inference rule is adequate, and if exceptions to the usual behavior are rare. Thus the nature and use of inference rules involved in decomposition strategies will be a promising topic for future research.

A second inference strategy, the use of the *availability heuristic* (Tversky & Kahneman, 1973), relies on the ease with which specific instances come to mind. The more easily an instance of the behavior and associated details

come to mind, the more recent or frequent the behavior appears. While recent or frequent events are indeed easier to recall, ease of recall is also influenced by other factors, such as the vividness or importance of a memory. Accordingly, the recency and frequency of vivid events is likely to be overestimated, whereas the recency and frequency of pallid events are underestimated, resulting in the frequent finding that rare and vivid events are overreported whereas pallid and mundane events are underreported (Bradburn et al., 1987).

A particularly important inference strategy was identified by Michael Ross and his collaborators, who found that respondents answer retrospective questions by using their present status with regard to the attribute under study as a benchmark, and invoke an implicit theory of self to assess whether their past standing on that attribute was similar to or different from their current status (for reviews, see Ross, 1988; Ross & Conway, 1986). With regard to many variables, people hold *implicit theories of stability and change*, often related to naive conceptions of life-span development, on which considerable interpersonal agreement has been documented (Ross, 1988). These theories allow them to infer their previous attitudes and behaviors by using their current attitude or behavior as an initial estimate, which they adjust according to their implicit theory. The resulting reports of previous attitudes and behaviors are correct to the extent that the implicit theory is accurate (see also Nisbett & Wilson, 1977).

So far, this approach has been tested primarily with retrospective reports of attitudes and opinions (e.g., Markus, 1986; Ross & Conway, 1986), but it is equally applicable to retrospective reports of behaviors and experiences. Frequently, individuals assume a rather high degree of stability, resulting in underestimates of the degree of change that has occurred over time. Accordingly, retrospective estimates of income (Withey, 1954) or of tobacco, marijuana, and alcohol consumption (Collins, Graham, Hansen, & Johnson, 1985) were found to be heavily influenced by respondents' income or consumption habits at the time of interview. On the other hand, when respondents have reason to believe in change, they will detect change, even though none occurred. For example, respondents who participated in a study of skills training (that did not improve their skills on any objective measure) subsequently reported that their skills were considerably poorer *before* they participated in the program. Presumably, they used their belief in the effectiveness of the training program to infer what their skills must have been before they "improved" (Conway & Ross, 1984). Similarly, participants in a pain treatment program were found to remember more pain than they had recorded during a baseline period, again reflecting their belief in program-induced change (Linton & Goteslam, 1983; Linton & Melin, 1982). As a final example, women's retrospective

reports of menstrual distress were found to be a function of their theory of the menstrual cycle: The more respondents believed that their menstrual cycles affected their well-being, the more their retrospective reports deviated from diary data obtained during the cycles (McFarland, Ross, & DeCourville, 1988).

Finally, respondents may use their general world knowledge to infer reasonable answers. Asked to report the frequency of dental visits, a respondent may refer to *normative expectations* (e.g., the expectation to have semi-annual checkups) and may adjust the resulting estimate to reflect individual deviations (e.g., "I don't go as often as I should.") (Bradburn et al., 1987, p. 160). This inference strategy is likely to result in estimates that are displaced toward the initial anchor supplied by the normative expectations, reflecting the general finding that initial anchor values dominate the resulting estimates (Tversky & Kahneman, 1974).

In general, these findings emphasize that retrospective behavioral reports are to a large degree theory driven: Respondents are likely to begin with some fragmented recall of the behavior under study and to apply various inference rules to arrive at a reasonable estimate. It will be an important task for future research to learn more about the kind of theories that respondents apply.

Editing the Answer

After having determined a private estimate of the frequency of the target behavior, the respondent has to report his or her estimate to the researcher. The communicated estimate may deviate from the respondent's private estimate due to considerations of social desirability and self-presentation. Practical steps taken to reduce response editing include various techniques to assure anonymity; wording the question in a way that implies that the undesirable behavior is "usual," thus decreasing respondents' concern (but potentially introducing other biases); using respondents' own words to label the target behavior; and embedding the question into a list of other, less sensitive ones. Sudman and Bradburn (1983, pp. 54ff.) provide a detailed discussion of various techniques employed by survey researchers.

USING RESPONSE ALTERNATIVES: WHAT RESPONDENTS LEARN FROM SCALES

The preceding discussion focused on how respondents report behavioral frequencies in response to open-ended questions, and emphasized the importance of various inference rules that respondents apply to compute

retrospective estimates. In line with Bradburn et al.'s (1987) observation that respondents will "use any information they have to generate a reasonable answer" (p. 160), the following sections of this chapter will explore the inference rules that respondents apply to information that is particularly salient in the research situation: namely, the information provided by the response alternatives presented to them by the researcher.

Frequently, respondents are asked to report their behavior by checking the appropriate alternative from a list of response categories provided to them. While the selected alternative is assumed to inform the researcher about the respondent's behavior, it is frequently overlooked that a given set of response alternatives may be far more than a simple "measurement device." Rather, it may also constitute a source of information for the respondent (see Schwarz, 1988; Schwarz & Hippler, 1987). Specifically, respondents were found to assume that the range of the response alternatives provided to them reflects the researcher's knowledge of, or expectations about, the distribution of the behavior in the "real world." To the extent that respondents apply this "naive theory" of response scales, the range of the response alternatives may affect respondents' understanding of the question, their behavioral reports, and subsequent related judgments. The following sections review a series of experiments that bear on these possibilities.

Understanding the Question

Frequently, the behavior under study is ill-defined and open to interpretation. This is particularly likely when researchers are interested in subjective experiences. Assume, for example, that respondents are asked to indicate how frequently they were "really irritated" recently. Before the respondent can give an answer, he or she must decide what the researcher means by "really irritated." Does this refer to major irritations such as fights with one's spouse or does it refer to minor irritations such as having to wait for service in a restaurant? If the respondent has no opportunity to ask the interviewer for clarification, or if a well-trained interviewer responds, "Whatever you feel is really irritating," he or she might pick up some pertinent information from the questionnaire. One such piece of information may be the frequency range provided by the scale.

For example, respondents who are asked to report how often they are irritated on a scale ranging from "several times daily" to "less than once a week" may relate the frequency range of the response alternatives to their general knowledge about the frequency of minor and major annoyances. Assuming that major annoyances are unlikely to occur "several times a day," they may

consider instances of less severe irritation to be the target of the question than may respondents who are presented a scale ranging from "several times a year" to "less than once every three months." Experimental data support this assumption (Schwarz, Strack, Mo, & Chassein, 1988). Respondents who reported their experiences on the former scale subsequently reported less extreme examples of annoying experiences than respondents who were given the latter scale. Thus the type of annoying experiences that respondents reported was determined by the frequency range of the response alternatives in combination with respondents' general knowledge, rather than by the wording of the question per se. Accordingly, the same question combined with different frequency scales is likely to assess different experiences.

Theoretically, the impact of the response alternatives on respondents' interpretation of the question should be more pronounced the less clearly the target behavior is defined. For this reason, questions about subjective experiences may be particularly sensitive to the impact of response alternatives because researchers usually refrain from providing a detailed definition of the target experience so as not to interfere with its subjective nature. Ironically, assessing the frequency of a behavior with precoded response alternatives may result in doing just what is avoided in the wording of the question.

Estimating Behavioral Frequencies

That a list of response alternatives may bias respondents' behavioral reports has repeatedly been observed by survey researchers (e.g., Bradburn, Sudman, & Associates, 1979) and a study on leisure time activities can serve as an illustration. In a study by Schwarz, Hippler, Deutsch, and Strack (1985, Experiment 1), a quota sample of German adults reported how many hours a day they spend watching TV. Previous research by Darschin and Frank (1982) indicated that the average daily TV consumption in West Germany was slightly more than 2 hours. To test the impact of different response alternatives, half of the sample received a scale ranging in half-hour steps from "up to ½ hour" to "more than 2½ hours," while the other half received a scale ranging from "up to 2½ hours" to "more than 4½ hours," as shown in Table 4.1.

The range of the response alternatives had a pronounced impact on respondents' reports. Specifically, only 16.2% of the respondents who were presented the low-frequency scale reported watching TV for more than 2½ hours, while 37.5% of the respondents who were presented the high-frequency scale did so.

TABLE 4.1 Reported Daily TV Consumption as a Function of Response Alternatives (N = 132)

<i>Low-Frequency Alternatives</i>	%	<i>High-Frequency Alternatives</i>	%
Up to ½ hour	7.4	Up to 2½ hours	62.5
½ hour to 1 hour	17.7	2½ hours to 3 hours	23.4
1 hour to 1½ hours	26.5	3 hours to 3½ hours	7.8
1½ hours to 2 hours	14.7	3½ hours to 4 hours	4.7
2 hours to 2½ hours	17.7	4 hours to 4½ hours	1.6
More than 2½ hours	16.2	More than 4½ hours	0

SOURCE: Adapted from Schwarz, Hippler, Deutsch, and Strack (1985).

Mediating processes. Two processes may contribute to this finding. On the one hand, the memory research reviewed above suggests that respondents may be unlikely to have detailed episodic memories of behaviors that are as frequent and mundane as watching TV. Rather, they may base their answers on salient information that allows the computation of a reasonable estimate. One source of pertinent information that is highly salient in the research context is the range of the response alternatives provided to them. Accordingly, respondents may use the range of the response alternatives as a frame of reference to estimate their own TV consumption.

As another theoretical possibility, respondents may be sensitive to self-presentational concerns when responding. They may be reluctant to check a response alternative that seems extreme in the context of the scale and thus reflects a presumably unusual behavior. This has been suggested by Bradburn and Danis (1984) in a discussion of higher reports of alcohol consumption in an open than in a closed response format (Bradburn et al., 1979).

Both hypotheses implicitly assume that respondents use the range of the response alternatives to infer which behavior is "usual." In general, respondents were found to assume that the behavior of the "average" person is represented by the values stated in the middle range of the response scale, and that the extremes of the scale also represent the extremes of the distribution—at least as long as these values do not appear obscure in the context of the respondents' lay theories (Schwarz et al., 1985; Schwarz & Hippler, 1987).

A number of experimental studies have investigated how these assumptions mediate the impact of response alternatives: Is the impact of response alternatives on behavioral reports mediated by self-presentation considera-

tions or by respondents' use of the range of the response alternatives as a frame of reference in estimating frequencies that are difficult to reconstruct from memory?

Self-reports versus proxy reports. One way to differentiate between the two proposed mechanisms is to compare the impact of response alternatives on reports of one's own behavior (self-reports) and on reports about the behavior of other persons (proxy reports). In general, the two process assumptions lead to opposite predictions for both types of reports. If the impact of response alternatives is mediated by self-presentation concerns, scale effects should be *stronger* when respondents report their own behavior than when they report the behavior of friends or distant acquaintances. This follows from the assumption that they are presumably more concerned about self-presentation than about the image they present of others.

If respondents use the values presented in the scale to compute an estimate, on the other hand, the impact of scale range should be more pronounced the less other information is available that could be used to compute an answer. Therefore, the effect of scale range should be *smaller* when respondents report their own behavior than when they report the behavior of friends or distant acquaintances, because they can draw upon a broader base of information that allows the reconstruction of relevant episodes for self-reports.

In one study, American undergraduates were asked to report their own weekly TV consumption, the weekly TV consumption of a close friend, or the weekly TV consumption of a "typical undergraduate" of their university on a scale ranging from "up to 2½ hours per week" to "more than 10 hours" or on a scale ranging from "up to 10 hours" to "more than 25 hours" (Schwarz & Bienias, in press, Experiment 1). As predicted by the frame of reference hypothesis, the impact of scale range was most pronounced when respondents estimated the TV consumption of a "typical undergraduate." Specifically, 71% provided estimates of more than 10 hours per week on the high-frequency response scale, but only 13% did so on the low-frequency scale, resulting in a difference of 58 percentage points. The impact of scale range was least pronounced, on the other hand, when respondents reported their own TV consumption, with a difference of 32 percentage points. This pattern of results is opposite to the one predicted by the self-presentation hypothesis, which holds that self-reports should be most strongly affected. Reports about the behavior of close friends fell between these extremes, as both hypotheses would predict, with a difference of 37 percentage points.

These findings suggest that respondents use the range of the response

alternatives as a frame of reference in estimating behavioral frequencies, and that the less other information they have, the more likely they are to rely on this frame. In this regard, it is informative to note that most of the respondents in the "friends" condition chose their roommates as target persons. It therefore comes as little surprise that their estimates of their friends' behaviors were only slightly more susceptible to scaling effects than their self-reports.

In summary, the impact of the range of the response alternatives increased as the availability of relevant information about the target behavior decreased. This conclusion is further supported by a study that used an individual difference approach to explore the impact of information accessibility and self-presentation concerns.

Private and public self-consciousness. Previous research in personality psychology has indicated that individuals who focus attention on the self provide more accurate self-reports, presumably because relevant self-knowledge is cognitively more accessible to them (see Wicklund, 1982, for a review). This suggests that these individuals should be less influenced by the ranges of the response scales provided to them because they may have better access to relevant self-related information. Such a finding would parallel the results of the previous study, further supporting the hypothesis that the impact of scale range decreases as respondents' available knowledge about the behavior under investigation increases.

However, individuals differ not only in the extent to which they pay attention to their own behaviors and feelings, but also in the extent to which they pay attention to the impression they give to others. According to the self-presentation hypothesis, individuals who care a lot about their public image should be more affected by scale range than individuals who pay less attention to what others think of them.

Accordingly, individuals' disposition to pay attention to what others think of them and their disposition to focus on their own behaviors and feelings were assessed with the "public" and "private self-consciousness" scales developed by Fenigstein, Scheier, and Buss (1975). Specifically, American college students reported their weekly TV consumption on one of the previously described high- or low-frequency response scales and their reports were analyzed as a function of their private and public self-consciousness scores (Schwarz & Bienias, in press, Experiment 3).

As predicted by the frame of reference hypothesis, the impact of scale range was more pronounced for respondents who scored low on private self-consciousness than for respondents who scored high on private self-consciousness. Specifically, 51% of the respondents who scored *low* on pri-

vate self-consciousness reported watching TV for more than 10 hours per week when given the high-frequency response scale, while only 13% of them did so when given the low-frequency scale, resulting in a difference of 38 percentage points. In contrast, respondents who scored *high* on the private self-consciousness scale were not significantly affected by the range of the response scales provided to them. This finding presumably reflects the higher accessibility of self-related information under high self-consciousness, and suggests that these respondents used information recalled from memory, rather than information provided by the scales, to estimate their TV consumption. To this extent, the present results parallel the findings of the self-reports versus proxy reports study by indicating that the impact of response alternatives decreases as the accessibility of other information increases. Finally, respondents' public self-consciousness—that is, their disposition to focus attention on what others think of them—did not affect the impact of the frequency range of the response scales.

Conclusion. In combination, these findings suggest that respondents use the range of the response alternatives provided to them as a frame of reference in estimating their own behavioral frequencies. Accordingly, the less other information is easily accessible in memory, the more pronounced is the impact of response alternatives on respondents' reports. Self-presentation concerns, on the other hand, do not seem to play a major role, at least not with nonthreatening questions such as the ones used here. It is conceivable, however, that self-presentation concerns that may be elicited by highly threatening questions will be compounded if the respondent discovers that his or her report requires the endorsement of a response alternative that seems extreme in the context of the list. Thus response alternatives may also affect behavioral reports at the editing stage of the judgmental process, although strong empirical evidence for this possibility has not yet been provided.

How strongly response alternatives bias respondents' reports will depend upon how much the scale deviates from respondents' actual behavioral frequencies. Theoretically, a precoded scale that matches respondents' actual behavioral frequencies may be expected to increase the validity of the obtained estimates. Note, however, that the effects of a given response scale may be different for different subpopulations. To the extent that the actual frequency of a behavior varies across subpopulations, a set of response alternatives constructed on the basis of pretest data is unlikely to match the actual behavior of extreme groups, which are likely to be underrepresented in pretests with limited sample sizes. Because the range of the scale may be used by all respondents as a frame of reference in estimating behavioral frequencies, it may therefore tend to obscure differences between subpopulations.

Comparative Judgments

If respondents use the range of response alternatives to make inferences about the distribution of the behavior in the population, we may also expect response alternatives to affect a wide range of related judgments to which these inferences may be relevant. Most important, checking one from an ordered set of response alternatives may be equivalent to determining one's own location in a distribution. For example, many of the respondents who reported their daily TV consumption on the low-frequency scale shown in Table 4.1, were likely to check a response category in the upper range of that scale. This may have suggested to them that they watch more TV than many other people. In contrast, respondents who received the high-frequency scale were likely to check a category in the lower range of that scale, which may have suggested to them that they watch less TV than is "usual." If so, the range of the response scale, and respondents' own placement on the scale, may provide comparison information, and respondents may use this information for subsequent judgments to which such information may be relevant. The range of the response alternatives therefore may not only determine the responses given to the particular question to which they pertain, but influence answers to subsequent questions as well.

To test these considerations, respondents of the leisure-time study described above (Schwarz et al., 1985, Experiment 1) were asked to evaluate how important a role TV plays in their own lives. As expected, respondents reported a higher importance of TV in their own lives when the low-frequency range suggested a low TV consumption to be typical than when the high-frequency range suggested that many people watch a lot of TV. Note that this was true even though the former respondents reported watching less TV than the latter, as discussed above.

In a related study, the range of the response scale used to report one's own TV consumption affected respondents' subsequent evaluation of their leisure time even under conditions where the crucial response scale and the leisure time question were separated by several buffer items (Schwarz et al., 1985, Experiment 2). In this study, respondents reported higher satisfaction with the variety of things they do in their leisure time when they had previously reported their own TV consumption on the high-frequency scale, which suggested to them that they watch less TV than average, than when they had given their report on the low-frequency scale. Thus it seems that the impact of the range of the response alternatives on subsequent comparative judgments is rather robust (see also Schwarz & Scheuring, *in press*).

Theoretically, the impact of the response alternatives on comparative

judgments should be more pronounced the less comparison information is easily available from other sources. In addition, researchers should be aware that precoded response alternatives may bias respondents' comparative judgments even under conditions where the set of response alternatives perfectly matches the actual distribution of behavior, thus introducing little bias in behavioral reports. While individuals who were not exposed to the response scale may use varying standards of comparison (e.g., based on their reference group), exposure to a given response scale may result in the use of the salient comparison information provided by the scale. To this extent, judgments obtained from a sample that was exposed to the response scale may differ from judgments prevailing in the population to which one may want to generalize the obtained results.

And What About the Users of a Respondent's Report?

However, respondents are not the only ones who use the range of response alternatives as a frame of reference in making comparative judgments. Rather, potential users of a respondent's report are also likely to be affected by the response scale in their interpretation of the reported behavior. In one study, experienced physicians and advanced students of medicine were more likely to assume that having a given physical symptom twice a week requires medical attention when this frequency was reported on a low- rather than a high-frequency range scale (Schwarz, Bless, Bohner, Harlacher, & Kellenbenz, 1988, Experiment 1). For example, practicing physicians were more likely to recommend that a fictitious patient should see a doctor if the patient reported vomiting twice a week on a low- rather than high-frequency range scale. Thus experienced professional users of symptom checklists evaluated the same frequency report differently, depending on the context provided by the frequency range of the response scale.

CONCLUSIONS

What are the implications of the reviewed findings for questionnaire construction? Most important, the reviewed findings demonstrate that respondents are unlikely to answer quantitative autobiographical questions on the basis of a "recall and count" procedure. While researchers may increase the likelihood that respondents will attempt to use a recall and count strategy by asking specific questions pertaining to a well-anchored short reference

period, respondents will usually base their answers on some fragmented recall from which they attempt to infer a plausible estimate using various inference strategies. In doing so, respondents "use any information they have to generate a reasonable answer" (Bradburn et al., 1987, p. 160). Accordingly, the traditional distinction between "opinion questions," which are presumably answered on the basis of somewhat unreliable judgmental processes, and "factual questions," which are presumably answered on the basis of more reliable recall from memory, is misleading. As much as answering opinion questions requires a considerable degree of recall (see, e.g., the discussions by Bodenhausen & Wyer, 1987; Schwarz & Strack, in press; Strack & Martin, 1987; Tourangeau & Rasinski, 1988), answering quantitative "factual" questions requires a considerable degree of inference and judgment. A judgmental approach to retrospective reports, emphasizing the inference and estimation strategies that underlie respondents' reports, may therefore prove to be an important supplement to current research on autobiographical memory, and is likely to improve our understanding of basic judgmental processes as well as the collection of retrospective data in theory testing and applied social and psychological research.

REFERENCES

- Baddeley, A. D., & Hitch, G. J. (1977). Recency reexamined. In S. Dornic (Ed.), *Attention and performance* (Vol. 6, pp. 647-667). Hillsdale, NJ: Erlbaum.
- Baddeley, A. D., Lewis, V., & Nimmo-Smith, J. (1978). When did you last . . . ? In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory* (pp. 77-83). London: Academic Press.
- 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.
- Blair, E., & Burton, S. (1987). Cognitive processes used by survey respondents to answer behavioral frequency questions. *Journal of Consumer Research*, 14, 280-288.
- Bodenhausen, G. V., & 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, & S. Sudman (Eds.), *Social information processing and survey methodology* (pp. 6-41). New York: Springer-Verlag.
- Bradburn, N., & Danis, C. (1984). Potential contributions of cognitive research to survey questionnaire design. In T. B. Jabine, M. L. Straf, J. M. Tanur, & R. Tourangeau (Eds.), *Cognitive aspects of survey methodology: Building a bridge between disciplines* (pp. 101-129). Washington, DC: National Academy Press.
- Bradburn, N. M., Rips, L. J., & Shevell, S. K. (1987). Answering autobiographical questions: The impact of memory and inference on surveys. *Science*, 236, 157-161.

- Bradburn, N., Sudman, S., & Associates (1979). *Improving interviewing methods and questionnaire design*. San Francisco: Jossey-Bass.
- Brown, N. R., Rips, L. J. & 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., & Shevell, S. K. (1986). Public memories and their personal context. In D. C. Rubin (Ed.), *Autobiographical memory* (pp.137-158). Cambridge: Cambridge University Press.
- Collins, L. M., Graham, J. W., Hansen, W. B., & Johnson, C. A. (1985). Agreement between retrospective accounts of substance use and earlier reported substance use. *Applied Psychological Measurement*, 9, 301-309.
- Conway, M., & Ross, M. (1984). Getting what you want by revising what you had. *Journal of Personality and Social Psychology*, 47, 738-748.
- Darschin, W., & Frank, B. (1982). Tendenzen im Zuschauerverhalten. Teleskopie-Ergebnisse zur Fernsehnutzung im Jahre 1981. *Media Perspektiven*, 4, 276-284.
- Fenigstein, A., Scheier, M. F., & Buss, A. H. (1975). Public and private self-consciousness: Assessment and theory. *Journal of Consulting and Clinical Psychology*, 43, 522-527.
- Groves, R. M., & Kahn, R. L. (1979). *Surveys by telephone: A national comparison with personal interviews*. New York: Academic Press.
- Gruneberg, M. M., Morris, P. E., & Sykes, R. N. (Eds.). (1988). *Practical aspects of memory: Current research and issues* (Vol. 1). Chichester: John Wiley.
- Hippler, H. J., & Schwarz, N. (1987). Response effects in surveys. In H. J. Hippler, N. Schwarz, & S. Sudman (Eds.), *Social information processing and survey methodology* (pp. 102-122). New York: Springer-Verlag.
- Hippler, H. J., Schwarz, N., & Sudman, S. (Eds.). (1987). *Social information processing and survey methodology*. New York: Springer-Verlag.
- Jabine, T. B., Straf, M. L., Tanur, J. M., & Tourangeau, R. (Eds.). (1984). *Cognitive aspects of survey methodology: Building a bridge between disciplines*. Washington, DC: National Academy Press.
- Linton, M. (1975). Memory for real-world events. In D. A. Norman & D. E. Remelhart (Eds.), *Explorations in cognition* (pp. 376-404). San Francisco: Freeman.
- Linton, M. (1982). Transformations of memory in everyday life. In U. Neisser (Ed.), *Memory observed: Remembering in natural contexts* (pp. 77-91). San Francisco: Freeman.
- Linton, S. J., & Goteslam, K. G. (1983). A clinical comparison of two pain scales: Correlation, remembering chronic pain and a measure of compliance. *Pain*, 17, 57-65.
- Linton, S. J., & Melin, L. (1982). The accuracy of remembering chronic pain. *Pain*, 13, 281-285.
- Loftus, E., & Fathi, D. C. (1985). Retrieving multiple autobiographical memories. *Social Cognition*, 3, 280-295.
- Loftus, E. F., & Marburger, W. (1983). Since the eruption of Mt. St. Helens, has anyone beaten you up? *Memory and Cognition*, 11, 114-120.
- Loftus, E. F., Smith, D. K., Johnson, D. A. & Fiedler, J. (1988). Remembering "when": Errors in the dating of autobiographical memories. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues* (Vol. 1, pp. 234-240). Chichester: John Wiley.
- Markus, G. B. (1986). Stability and change in political attitudes: Observed, recalled, and explained. *Political Behavior*, 8, 21-44.

- Mathiowetz, N. A. (1986, June). *Episodic recall and estimation: Applicability of cognitive theories to survey data*. Paper presented at the Social Science Research Council Seminar on Retrospective Data, New York.
- McFarland, C., Ross, M., & DeCourville, N. (1988). *Retrospective vs. current reports of menstrual distress*. Unpublished manuscript, University of Waterloo.
- Means, B., Mingay, D. J., Nigam, A., & Zarrow, M. (1988). A cognitive approach to enhancing health survey reports of medical visits. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues* (Vol. 1, pp. 537-542). Chichester: John Wiley.
- Neisser, U. (Ed.). (1982). *Memory observed*. San Francisco: Freeman.
- Neisser, U. (1986). Nested structure in autobiographical memory. In D. C. Rubin (Ed.), *Autobiographical memory* (pp. 71-88). Cambridge: Cambridge University Press.
- Neter, J., & Waksberg, J. (1964). A study of response errors in expenditure data from household interviews. *Journal of the American Statistical Association*, 59, 18-55.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231-259.
- Reiser, B. J., Black, J. B., & Abelson, R. P. (1985). Knowledge structure in the organization and retrieval of autobiographical memories. *Cognitive Psychology*, 17, 89-137.
- Ross, M. (1988). *The relation of implicit theories to the construction of personal histories*. Unpublished manuscript, University of Waterloo.
- Ross, M., & Conway, M. (1986). Remembering one's own past: The construction of personal histories. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition* (pp. 122-144). New York: Guilford.
- Rubin, D. C. (Ed.). (1986). *Autobiographical memory*. Cambridge: Cambridge University Press.
- Schwarz, N. (1987, April). *Cognitive aspects of labor surveys in a multi-national context* (OECD Document MAS/WP 7[87]4). Report commissioned for the ninth meeting of the Working Party on Labor Statistics. Organization for Economic Cooperation and Development, Paris.
- Schwarz, N. (1988). Was Befragte aus Antwortalternativen lernen. *Planung und Analysen*, 15, 103-107.
- Schwarz, N., & Bicnias, J. (in press). What mediates the impact of response alternatives on frequency reports of mundane behaviors? *Applied Cognitive Psychology*.
- Schwarz, N., Bless, H., Bohner, G., Harlacher, U., & Kellenbenz, M. (1988). Response scales as frames of reference: The impact of frequency range on diagnostic judgments. *ZUMA-Arbeitsbericht*, 6.
- Schwarz, N., & Hippler, H. J. (1987). What response scales may tell your respondents. In H. J. Hippler, N. Schwarz, & S. Sudman (Eds.), *Social information processing and survey methodology* (pp. 163-178). New York: Springer-Verlag.
- Schwarz, N., Hippler, H. J., Deutsch, B., & Strack, F. (1985). Response categories: Effects on behavioral reports and comparative judgments. *Public Opinion Quarterly*, 49, 388-395.
- Schwarz, N., & Scheuring, B. (in press). Judgments of relationship satisfaction: Inter- and intraindividual comparisons as a function of questionnaire structure. *European Journal of Social Psychology*.
- Schwarz, N., & Strack, F. (in press). Evaluating one's life: A judgment model of subject-

- live well-being. In F. Strack, M. Argyle, & N. Schwarz (Eds.), *The social psychology of well-being*. London: Pergamon.
- Schwarz, N., Strack, F., Mo, G., & Chassin, B. (1988). The range of the response alternatives may determine the meaning of the question. *Social Cognition*, 6, 107-117.
- Strack, F., & Martin, L. (1987). Thinking, judging, and communicating: A process account of context effects in attitude surveys. In H. J. Hippler, N. Schwarz, & S. Sudman (Eds.), *Social information processing and survey methodology* (pp. 123-148). New York: Springer-Verlag.
- Strube, G. (1987). Answering survey questions: The role of memory. In H. J. Hippler, N. Schwarz, & S. Sudman (Eds.), *Social information processing and survey methodology* (pp. 86-101). New York: Springer-Verlag.
- Sudman, S., & Bradburn, N.M. (1983). *Asking questions*. San Francisco: Jossey-Bass.
- Sudman, S., Finn, A., & Lannon, L. (1984). The use of bounded recall procedures in single interviews. *Public Opinion Quarterly*, 48, 520-524.
- Sudman, S., & Schwarz, N. (1987). Contributions of cognitive psychology to marketing research. *ZUMA-Arbeitsbericht* 9.
- Thompson, C.P. (1982). Memory for unique personal events: The roommate study. *Memory and Cognition*, 10, 324-332.
- Tourangeau, R., & Rasinski, K. A. (1988). Cognitive processes underlying context effects in attitude measurement. *Psychological Bulletin*, 103, 299-314.
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5, 207-232.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 85, 1124-1131.
- Wagenaar, W. A. (1986). My memory: A study of autobiographical memory over six years. *Cognitive Psychology*, 18, 225-252.
- Wagenaar, W. A. (1988). People and places in my memory: A study on cue specificity and retrieval from autobiographical memory. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues* (Vol. 1, pp. 228-232). Chichester: John Wiley.
- Whitten, W. B., & Leonard, J. M. (1981). Directed search through autobiographical memory. *Memory and Cognition*, 9, 566-579.
- Wickelgren, W. A. (1976). Memory storage dynamics. In W. K. Estes (Ed.), *Handbook of learning and cognition processes* (Vol. 4, pp. 321-362). Hillsdale, NJ: Erlbaum.
- Wicklund, R. A. (1982). Self-focused attention and the validity of self-reports. In M. P. Zanna, E. T. Higgins, & C. P. Herman (Eds.), *Consistency in social behavior* (pp. 149-172). Hillsdale, NJ: Erlbaum.
- Williams, M. D., & Hollan, J. D. (1981). The process of retrieval from very long term memory. *Cognitive Science*, 5, 87-119.
- Withey, S. B. (1954). Reliability of recall of income. *Public Opinion Quarterly*, 18, 31-34.