Journal of Official Statistics, Vol. 11. No. 1, 1995, pp. 79–92

A Decade of Questions

Nora Cate Schaeffer[1]

1. Introduction

During *JOS*'s first decade, the set of issues survey methodologists and other researchers recognize as important in writing survey questions and designing survey instruments became increasingly diverse. Research about survey questions and instruments now draws on or contributes to a wide variety of disciplines and subdisciplines including cognitive psychology, conversation analysis, ethnomethodology, cognitive anthropology, discourse analysis, pragmatics, and computer science. Although this review will focus on developments in the last decade, it would still be impossible to summarize systematically ten years of research relevant to instrument design in as many pages. So instead I will selectively characterize some of the directions recent work has taken, point to research rather than summarize it, and note a few issues that future inquiry faces.

The purpose for which survey questions and instruments are developed provides the context for this discussion. Survey questions and instruments are used as part of measurement operations in surveys that are conducted to describe a group or population. The adequacy of survey measurement operations is ultimately evaluated using some criterion – for example, compatibility with an established theory, psychometric reliability, agreement with administrative records, and so on. The constructs we attempt to measure – whether behaviors, events, feelings, attitudes, other judgments, or social characteristics – are knowable only through some measurement procedure. Thus, our observations are necessarily "contaminated" by the process of measurement, and all scientific descriptions and tests are to some extent conditional on the method of observations.

The themes I consider within this context include the application of theories and techniques imported from cognitive psychology to the design of survey questions, the study of interaction in the interview and its relation to question design, a revisiting of the debate about standardization, the development and systemization of methods for testing survey questions, the application of more powerful analytic techniques to the study of measures of subjective phenomena, and the effect of new technologies for administering survey instruments. In some cases, these developments have different implications for researchers who are primarily interested in questions about behaviors and events (many of them working within government statistical agencies) and researchers whose focus is questions about other subjective phenomena (most of them probably working in universities). Most of this research is concerned with

[1] University of Wisconsin, Department of Sociology, 1180 Observatory Drive, Madison, WI 53706, U.S.A.
Acknowledgements: I would like to thank Elizabeth Martin, Stanley Presser, and Eleanor Singer for their helpful suggestions at various stages in the preparation of this manuscript.

instruments designed to be used by an interviewer, but some attention has also been given to self-administered instruments.

2. Cognitive Psychology and Survey Questions

In the United States, those working on the redesign of the National Crime Survey and the Panel on Survey Measurement of Subjective Phenomena of the Committee on National Statistics were early in emphasizing the potential importance of insights from cognitive psychology for improving question design (see Biderman and Moore 1980; Turner and Martin 1984; the review of NCS work in Martin 1993; and the history summarized in Jobe and Mingay 1991). The programmatic statement *Cognitive Aspects of Survey Methodology* (Jabine, Straf, Tanur, and Tourangeau 1984) offered vivid and persuasive suggestions of how such cross-fertilization might improve survey questions and the data they obtain. Cross-disciplinary collaborations are probably most durable and productive when both disciplines benefit (see Jobe, Tourangeau, and Smith 1993), and survey interviews using probability samples offer a naturally occurring applied setting for testing many theories about cognition and memory.[2]

Research drawing on the techniques and theories of cognitive psychology has considered a wide range of constructs including both events and behaviors – such as dental visits (Lessler, Tourangeau, and Salter 1989), medical events (e.g., Means, Nigam, Zarrow, Loftus, and Donaldson 1989; Means and Loftus 1991), dietary intake (Smith 1991), smoking (Means, Swan, Jobe, and Esposito 1992), and cancer screening procedures (Sudman, Warneke, Johnson, O'Rourke, and Davis 1994) – and various subjective phenomena, such as attitudes (e.g., Krosnick 1991; Krosnick and Alwin 1987; Tourangeau 1992; Tourangeau, Rasinski, and Bradburn 1991). Recent edited volumes consider the application of social-information-processing models to surveys (Hippler, Schwarz, and Sudman 1987); context effects, primarily for questions about subjective phenomena (Schwarz and Sudman 1992); the accuracy of memories of events (Schwarz and Sudman 1994); methods for determining cognitive processes in answering questions (Schwarz and Sudman, forthcoming); and wide-ranging "inquiries into the cognitive bases of surveys" (Tanur 1992). In addition, there have been general explorations of cognitive processes, such as recall strategies and telescoping, that have important implications for survey measurement (e.g., Blair and Burton 1987; Huttenlocher, Hedges, and Bradburn 1990; Menon 1993). For those concerned with self-administered instruments, cognitive psychology has provided both theories (e.g., Strack and Schwarz 1992) and methods for solving practical design problems (e.g., Jenkins and Dillman 1993).[3] Taken together, this research sug-

[2] Jobe and Mingay's (1991, p. 182) comment that surveys provide "ecological validity" for the work of cognitive psychologists is a nice contrast with Cicourel's (1982) argument that surveys lack ecological validity because the social context of the survey interview is remote from daily life. These contrasting perspectives are also relevant, however, to the wider interests of survey methodologists, who are ultimately interested in the quality of survey data as measurements of constructs.

[3] Research conducted in preparation for the 2000 census has resulted in other improvements in designing self-administered instruments and implementing mail surveys (e.g., Dillman, Clark, and Treat 1994; Dillman, Sinclair, and Clark 1993).

gests that there could be science to the art of asking questions and begins to build a foundation for such theories. We could have, if not a single theory to guide the writing of survey questions, at least theories to use in this task.

The research just mentioned also has implications for the measurement enterprise of which survey questions are a component. A review of, for example, studies in which self-reports of health care events are compared with providers' records (e.g., Jobe et al. 1990 and Loftus, Smith, Klinger, and Fiedler 1992; for earlier studies see, e.g., Andersen, Kasper, Frankel 1979; Cannell 1965) suggests that the accuracy of such self-reports would probably be limited even with well designed questions (and interviewers capable of administering them), because of the constraints of interview time, the respondent's motivation, the respondent's memory (both encoding and recall), the interviewer's skill, and so on. Even if one accepts that instrument design cannot completely eliminate response errors, one might at least hope that it could help make the structure of errors similar across respondents. However, this result may also be beyond question design; for example, respondents with a complex history of medical visits may report less accurately than respondents who made no medical visits, no matter how a question sequence is written. Such considerations reinforce the importance of considering the role of a survey question within the larger task of measurement: Knowledge that improves "question design" is partly – though far from entirely – knowledge of how features of questions inescapably contribute "invalid" variance to answers; this understanding in turn increases our ability to control, model, or interpret features of the measurement process (e.g, Alwin 1991). In addition to their contributions to improved measurement, however, advances in question design are also inextricably linked to advances in understanding of the thing being measured and contribute to that understanding.

Much of the research just referred to focuses on four stages of an information-processing model of the response process: the comprehension of the question, the retrieval or construction of an answer, an evaluation of or judgment about the answer, and the reporting of the answer (Cannell, Marquis, and Laurent 1977; Cannell, Miller, and Oksenberg 1981; see also Tourangeau 1984). This model describes what goes on inside the head of the respondent when she answers a survey question. In contrast, the earlier model of Sudman and Bradburn (1974) gave more consideration to the social configuration of the survey interview: interviewer, respondent, and task. Think-aloud protocols, in which a respondent is asked to "think out loud" while answering a question in order to display his cognitive processing, complement traditional split-ballot experiments, which emphasize task characteristics, in the same way that the information-processing models complement the earlier model of situational influences on responses. Recent attempts to extend cognitive-processing models to consider how the interviewer and respondent together complete the interview can be seen as an attempt to unite cognitive-processing and structural frameworks for studying the interview (Esposito and Jobe 1991; Mullin, Conrad, Sander, and Herrmann 1993).

As attempts to integrate these two views of the survey interview illustrate, it is important that research on question design not lose sight of the social component of survey measurement. Measurement is "social" in at least three ways that are

particularly relevant to the design of survey questions:[4] It is socially structured experience that gives rise to memories or internal conditions (such as attitudes) that are either easy or difficult to report accurately (see, for example, Mathiowetz and Duncan 1988; Rodgers, Brown, and Duncan 1993; Schaeffer 1994; Krosnick 1991). In addition, the cognitive processing of the respondent takes place jointly with the cognitive processing of the interviewer, in the medium of interaction, so that "cognitive processing" may be indistinguishable from "interaction" (e.g., Schaeffer and Maynard forthcoming). And, finally, this interaction requires and draws on a socially organized substrate of interactional skills and conventions of talk, so that survey measurement is embedded in and conditioned on social processes (Krosnick, Li, and Lehman 1990; Schwarz 1994; review in Schaeffer 1991).

3. Interaction in the Survey Interview

Survey instruments are ultimately enacted. In an interview, the respondent experiences the question not as it appears on the page, but as it is delivered by the interviewer. To improve survey instruments, at some point we must look at what actually happens in the interview (or between the respondent and the self-administered instrument). Interaction is an inherently unwieldy subject of study – which is one reason we use standardized instruments to collect data. Early studies of interaction in the interview (e.g., Cannell, Fowler, and Marquis 1968; Marquis 1971) led to systems for coding features of the interaction that found applications in monitoring inverviewers (Cannell, Lawson, and Hausser 1975) and in testing questions (see below). Building on these studies and those of Brenner (1981, 1982), one body of recent work on interaction experimentally manipulates the behavior of interviewers to examine the consequences of different interviewing styles for behavior and the quality of the data (e.g., Dijkstra 1987; van der Zouwen, Dijkstra, and Smit 1991). Another set of more exploratory studies examines interaction in the interview in closer detail to discover issues relevant for designing and testing survey questions (e.g., Houtkoop 1994; Schaeffer and Maynard 1995; Cradock, Maynard, and Schaeffer 1993). For example, although previous studies indicate that respondents' direct requests for clarification might be relatively rare (e.g., Oksenberg, Cannell, and Kalton 1991), studies that examine the details of interaction find that respondents sometimes solicit help from the interviewer by a tentative reporting of potentially relevant facts. Thus, some responses that might be classified as "inadequate answers" during behavior coding are probably indirect requests for clarification.

Both types of study of interaction in the interview also potentially offer insights for those concerned about the training of interviewers, a topic which has received too little study (but see Fowler and Mangione 1990). Studies of interaction can reveal situations that interviewers are not prepared for and document how features of the

[4] The chapters on "The Role of the Respondent" and "The Question-and-Answer Process" in *Surveying Subjective Phenomena* (Turner and Martin 1984) consider the social context within which the respondent's cognitive processing takes place more fully than do most formulations; these chapters offer many useful ideas that research has yet to explore fully. The two volumes of *Surveying Subjective Phenomena* also offer one of the most extensive considerations available of the role of survey questions as part of the process of measurement.

instrument and interaction undermine standardization. Finally, a more thorough knowledge of how interviewers and respondents actually behave under the current regime of standardized interviewing may help us to conceive of alternative methods of standardized interviewing that, for some topics, might provide many of the benefits of the current method of standardization without some of its costs.

4. Standardization and its Discontents

Renewed discussion of the problems engendered and encountered by standardized interviewing (Briggs 1986; Mishler 1986; Suchman and Jordan 1990) raises useful challenges for question and instrument design. Criticisms of traditional standardized interviewing are particularly effective when taken together with research which suggests that the recall of events may be improved by procedures that do not fit neatly within the linear structure of standardized interviews (e.g., Means, Swan, Jobe, and Esposito 1992), that a less formal style of standarized interviewing may be more motivating to respondents than traditional formal standardized interviewing (e.g., Dijkstra 1987), and that interviewers do not always implement standardization well. The challenge for those designing survey questions is to weigh these problems against the well-known danger of errors introduced by interviewers; the capabilities of interviewing staffs; the different constraints entailed by the measurement of events and behaviors, subjective phenomena, or social characteristics; the limitations imposed by the interviewing technology used (see below); and the requirements of obtaining data to describe a population. Many failures of standardized interviewing are due to poor instrumentation (Collins 1980 provides striking examples), but even when it is well designed, a formal standardized interview may not be the best social environment for stimulating and motivating recall of complex topics.

5. Methods for Testing Survey Questions

Question testing has expanded beyond the "traditional" method, a small-scale pretest accompanied by an interviewer debriefing, to include the review of standardized questions by focus groups (e.g., Joseph et al. 1984), the use of cognitive laboratory techniques such as think-aloud protocols and paraphrasing (see, for example, Belson 1986; Lessler, Tourangeau, and Salter 1989), the application of cognitive techniques such as comprehension probes and vignettes (Fowler 1992; Oksenberg et al. 1991; Martin, Campanelli, and Fay 1991), respondent debriefing (part of the procedure for developing new labor force questions for the Current Population Survey described by Polivka and Rothgeb 1993; Campanelli, Martin, and Rothgeb 1991, Campanelli, Martin, and Creighton 1989), expert review of questions (Forsyth, Lessler, and Hubbard 1992), and behavior coding of pretest or pilot interviews (e.g., Oksenberg et al. 1991). (See also DeMaio et al. 1993 and the overview of methods in Forsyth and Lessler 1991.) Any technique that secures more time and budgetary resources for testing an instrument and trying it out with people different from the investigators is almost sure to improve the instrument. But this proliferation of methods invites an assessment of their comparative strengths which has barely begun (but see Presser and Blair 1994; Bischoping 1989) and which immediately faces the problem of what criterion to use in the assessment.

There are also many outstanding questions about both the implementation of these methods (e.g., the consequences of different ways of presenting a think-aloud procedure to a respondent) and the way we analyze and draw conclusions from the data they provide. Some questions of implementation could be answered by accumulated experience, if that experience were recorded systematically. We also have few tested methods for the many substantive and design decisions that are part of the question development process that precedes testing, and the literature supplies few models. One of these is Schuman and Presser's (1981) description of a multi-step procedure for the development of a set of categories for closed questions about qualities valued in a job and in children. (Schaeffer and Thomson 1992 give a different kind of illustration and references to still others.)

6. Subjective Phenomena and Measurement Models

Both our theoretical understanding of attitudes and other subjective phenomena and the methods we use to assess them have benefited from explorations of how self-reports are constructed and of how they are affected by the structure of a survey question (see sources cited above; also, e.g., Krosnick and Schuman 1988; Bishop, Oldendick, and Tuchfarber 1984). In addition, however, more powerful analytic techniques and more comprehensive data have been used to revisit problems such as assessing the effect on reliability of the number of categories in a rating scale, the category labels, or the presence of a "don't know" category (Alwin 1992; Alwin and Krosnick 1991; McClendon and Alwin 1993) and the consequences of controlling for acquiescence in analysis (Watson 1992; see also McClendon 1991). These studies help resolve issues that commonly arise in designing questions, particularly when the analysis examines how features of question structure (such as the number and labeling of categories) interact. But the ultimate usefulness of studies of the effects of item characteristics depends on the strength of the underlying analysis of question structure. We still lack a comprehensive, theoretically grounded classification of questions to guide such research (but see Appendix H of Turner and Martin 1984 for one attempt with a slightly different focus).

7. Technology and Instrumentation

Although the implications of computer-assisted interviewing (CAI) methods for instrument design have received some discussion (e.g., House and Nicholls 1988; Groves et al. 1988; Saris 1991), much of the discussion has concerned issues such as computer "fills" of variable information in question wording.[5] It is probably close to an unqualified benefit to have the computer "fill," for example, the name of the household member a question refers to, instead of having the interviewer insert the name as she reads the question (unless the fills are so numerous and complex that they slow the computer system). But other capabilities of computer-assisted inter-

[5] This discussion is limited to computer-assisted telephone interviewing (CATI), personal interviewing (CAPI), and "self-interviewing" (CASI). Instrumentation issues for other computer-assisted methods (such as prepared data entry) are not considered. I also do not consider related issues such as interviewer's acceptance of CAPI (Couper and Burt 1994).

viewing systems raise questions. CAI methods also allow the interviewer to present inconsistent answers or unusual answers to the respondent for resolution; but we do not know the consequences of the question design choices involved. A closely related feature of computer-assisted interviewing systems is more problematic: the ability to use information the respondent provided in a previous interview as part of "dependent interviewing" about change. There are many possible ways to present an answer a respondent gave in a previous interview in a question that asks for an update, and these different approaches could establish very different "thresholds" for reporting change or different degrees of pressure for consistency. The resulting thresholds and pressures could be thought of as creating "context" effects. As longitudinal surveys adopt computer-assisted interviewing methods, a variety of largely untested designs for conducting dependent interviewing is likely to be adopted.

Some implications of computer-assisted technologies for instrument design are quite broad and include effects on the timing, rhythm, and flexibility of initial design work, testing, and documentation (for some issues, see Dibbs and Hale 1993). Innovations such as computer-assisted self-interviewing in which the respondent hears questions read over headphones ("audio-CASI") make the privacy of a self-administered form available to respondents who may have difficulty reading, and may increase the accuracy of answers about sensitive topics (e.g., O'Reilly et al. 1994). Technology also has consequences for question design considered more narrowly. For example, the size and graphical limitations of the computer screen will constrain question design for some years to come. Some complex question formats – such as grids or calendars – can be translated to small screens, but there has been little assessment of what is lost in the process or of the details of how interviewers manage these complex forms.[6] Some consequences of the "linearity" of computer-assisted methods, such as the "segmentation" effect – which describes the disorientation of interviewers who see only "segments" of the instrument at any one time – are more problematic. To overcome this effect, some CATI and CAPI systems provide the interviewer with summaries, outlines, and other aids; designing these effectively requires attention to the "cognitive processing" of interviewers. As computer-assisted instruments become more widespread, the segmented linear character of the computer-assisted instrument may also limit how open or flexible "standardized" interviewing could become. For example, Suchman and Jordan (1990) suggest that the interviewer and respondent both review schedule, presumably in part so that the respondent could obtain an overview of the task and its objective – an experience that a CAPI instrument could not duplicate, at least not in the same way. Research like that reviewed earlier could end up suggesting that a respondent remembers more accurately if he can report relevant details in any order, view the details recalled so far to stimulate additional recall, and revise answers as needed as memories improve, but such an instrument could be difficult to implement on current computer-assisted interviewing systems.

[6] There is also little such research about paper instruments (but see Sanchez 1992; Freedman et al. 1988).

8. Conclusion

This brief overview is necessarily selective. It overlooks, for example, recent work about determining household structure and social characteristics (e.g., Martin and Griffin 1994; Gerber and Bates 1994; Kearney, Tourangeau, Shapiro, and Ernst 1993; Martin, DeMaio, and Campanelli 1990), studies of a wide range of specific question features such as response alternatives (e.g., Wright, Gaskell, and O'Muircheartaigh 1994; Schaeffer 1992), and related work such as that on mode of administration effects (e.g., Aquilino 1994). Even so, the studies considered here suggest that researchers working on issues in instrument design have taken seriously the need to develop the theoretical grounding for their work (e.g., Groves 1987) and have been inventive and wide-ranging in their search for ways of improving survey measurement.

9. References

Alwin, D. (1991). Research on Survey Quality. Sociological Methods and Research, 20, 3–29.

Alwin, D. (1992). Information Transmission in the Survey Interview: Number of Response Categories and the Reliability of Attitude Measurement. In Sociological Methodology, ed. P. Marsden. Washington, D.C.: The American Statistical Association.

Alwin, D.F. and Krosnick, J.A. (1991). The Reliability of Survey Attitude Measurement: The Influence of Question and Respondent Attributes. Sociological Methods and Research, 20, 139–181.

Andersen, R., Kasper, J., and Frankel, M.R. (1979). Total Survey Error: Applications to Improve Health Surveys. San Francisco: Jossey Bass.

Aquilino, W.S. (1994). Interview Mode Effects in Surveys of Drug and Alcohol Use: A Field Experiment. Public Opinion Quarterly, 58, 210–240.

Belson, W.A. (1986). Validity in Survey Research. Aldershot, England: Gower.

Biderman, A.D. and Moore, J.C. (1980). Report of the Workshop on Applying Cognitive Psychology to Recall Problems of the National Crime Survey. Unpublished manuscript. Washington, D.C.: Bureau of Social Science Research.

Bischoping, K. (1989). An Evaluation of Interviewer Debriefing in Survey Pretests. In New Techniques for Pretesting Survey Questions: Final report, eds. C. Cannell, L. Oksenberg, G. Kalton, K. Bischoping, and F. Fowler. Report supported by Grant number HS 05616 from the National Center for Health Services Research and Health Care Technology Assessment.

Bishop, G.F., Oldendick, R.W., and Tuchfarber, A. (1984). What Must My Interest in Politics Be If I Just Told You 'I Don't Know'? Public Opinion Quarterly, 48, 510–524.

Blair, E. and Burton, S. (1987). Cognitive Processes Used by Survey Respondents to Answer Behavioral Frequency Questions. Journal of Consumer Research, 14, 280–288.

Brenner, M. (1981). Patterns of Social Structure in the Research Interview. In Social Method and Social Life, ed., M. Brenner. New York: Academic Press.

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

Briggs, C.L. (1986). Learning How to Ask. New York: Cambridge University Press.

Campanelli, P.C., Martin, E.A., and Creighton, K.P. (1989). Respondents' Understanding of Labor Force Concepts: Insights from Debriefing Studies. Proceedings of the U.S. Bureau of the Census 1989 Annual Research Conference, 361–391.

Campanelli, P.C., Martin, E.A., and Rothgeb, J.M. (1991). The Use of Respondent and Interviewer Debriefing Studies as a Way to Study Response Error in Survey Data. The Statistician, 40, 253–264.

Cannell, C.F. (1965). Reporting of Hospitalization in the Health Interview Survey. National Center for Health Statistics, Series 2, No. 6, Washington, D.C.: Government Printing Office.

Cannell, C.F., Fowler, F.J., and Marquis, K.H. (1968). The Influence of Interviewer and Respondent Psychological and Behavioral Variables on the Reporting in Household Interviews. Vital and Health Statistics, Series 2, no. 26.

Cannell, C.F., Lawson, S.A., and Hausser, D.L. (1975). A Technique for Evaluating Interviewer Performance. Ann Arbor, MI: Survey Research Center, The University of Michigan.

Cannell, C.F., Marquis, K., and Laurent, A. (1977). A Summary of Studies of Interviewing Methodology: 1959–1970. Vital and Health Statistics, series 2, no. 69.

Cannell, C.F., Miller, P., and Oksenberg, L. (1981). Research on Interviewing Techniques. Sociological Methodology 1981, ed. S. Leinhardt. San Francisco: Jossey Bass.

Cicourel, A.V. (1982). Interviews, Surveys, and the Problem of Ecological Validity. American Sociologist, 17, 11–20.

Collins, M. (1980). Interviewer Variability: A Review of the Problem. Journal of the Market Research Society, 22, 77–95.

Couper, M.P. and Burt, G. (1994). Interviewer Attitudes Toward Computer-Assisted Personal Interviewing (CAPI). Social Science Computer Review, 12, 38–54.

Cradock, R.M., Maynard, D.W., and Schaeffer, N.C. (1993). Re-Opening Closed Answers: Managing Respondents' Elaborations on Categorical Answers in Standardized Interviews. University of Wisconsin-Madison, Center for Demography and Ecology, Working Paper 93–24.

DeMaio, T., Mathiowetz, N., Rothgeb, J., Beach, M.E., and Durant, S. (1993). Protocol for Pretesting Demographic Surveys at the Census Bureau. U.S. Bureau of the Census.

Dibbs, R. and Hale, A. (1993). Questionnaire Design in a Paperless Society. Proceedings of the U.S. Bureau of the Census 1993 Annual Research Conference, Arlington, VA, March, 353–361.

Dijkstra, W. (1987). Interviewing Style and Respondent Behavior: An Experimental Study of the Survey Interview. Sociological Methods and Research, 16, 309–334.

Dillman, D.A., Clark, J.R., and Treat, J.B. (1994). Influence of 13 Design Factors on Completion Rates to Decennial Census Questionnaires. Proceeding of the U.S. Bureau of the Census 1994 Annual Research Conference, Arlington, VA, March.

Dillman, D.A., Sinclair, M.D., and Clark, J.R. (1993). Response to Occupant-Addressed Census Surveys. Public Opinion Quarterly, 57, 289–304.

Esposito, J. and Jobe, J.B. (1991). A General Model of the Survey Interaction Process. Proceedings of the U.S. Bureau of the Census 1991 Annual Research Conference, Arlington, VA, March, 537–560.

Forsyth, B.H. and Lessler, J.T. (1991). Cognitive Laboratory Methods: A Taxonomy. In Measurement Errors in Surveys, eds. P. Biemer, R. Groves, L. Lyberg, N. Mathiowetz, and S. Sudman. New York: John Wiley.

Forsyth, B.H., Lessler, J. and Hubbard, M. (1992). Cognitive Evaluation of the Questionnaire. In Survey Measurement of Drug Use: Methodological Studies, eds. C.F. Turner, J.T. Lessler, and Joseph C. Gfroerer, U.S. Department of Health and Human Resources.

Fowler, F.J. (1992). How Unclear Terms Affect Survey Data. Public Opinion Quarterly, 56, 218–231.

Fowler, F.J. and Mangione, T.W. (1990). Standardized Survey Interviewing: Minimizing Interviewer-Related Error. Newbury Park: Sage.

Freedman, D., Thornton, A., Camburn, D., Alwin, D., and Young-DeMarco, L. (1988). The Life History Calendar: A Technique for Collecting Retrospective Data. In Sociological Methodology 1988, ed. C. Clogg, Washington, D.C.: the American Sociological Association.

Gerber, E.R. and Bates, N. (1994). Respondents' Understanding of Residence Terminology: Cognitive Research in the Living Situation Survey. Proceedings of the U.S. Bureau of the Census 1994 Annual Research Conference, Arlington, VA, March.

Groves, R.M. (1987). Research on Survey Data Quality. Public Opinion Quarterly, 51, S156–172.

Groves, R.M., Biemer, P.P., Lyberg, L.E., Massey, J.T. and Nicholls II, W.L., and Waksberg, J. eds. (1988). Telephone Survey Methodology. New York: John Wiley.

Hippler, H.-J., Schwarz, N., and Sudman, S. eds. (1987). Social Information Processing and Survey Methodology. New York: Springer Verlag.

House, C.C. and Nicholls II, W.L. (1988). Questionnaire Design for CATI: Design Objectives and Methods. In Telephone Survey Methodology, ed., R. Groves et al., New York: John Wiley.

Houtkoop, H. (1994). Meeting Both Ends: Between Standardization and Recipient Design in Telephone Survey Interviews. In Situated Order, eds. P. Ten Have and G. Psathas. Lanham, Maryland: University Press of America.

Huttenlocher, J., Hedges, L.V., and Bradburn, N.M. (1990). Reports of Elapsed Time: Bounding and Rounding Processes in Estimation. Journal of Experimental Psychology: Learning, Memory, and Cognition, 16, 196–213.

Jabine, T., Straf, M., Tanur, J.M., and Tourangeau, R. eds. (1984). Cognitive Aspects of Survey Methodology: Building a Bridge Between Disciplines. Washington, D.C.: National Academy Press.

Jenkins, C. and Dillman, D.A. (1993). Combining Cognitive and Motivational Research Perspectives for the Design of Respondent-Friendly Self-Administered Questionnaires. Paper presented at the annual meetings of the American Association for Public Opinion Research, St. Charles, IL, May.

Jobe, J.B., Tourangeau, R., and Smith, A.F. (1993). Contributions of Survey Research to the Understanding of Memory. Applied Cognitive Psychology, 7, 567–584.

Jobe, J.B. and Mingay, D.J. (1991). Cognition and Survey Measurement: History and Overview. Applied Cognitive Psychology, 5, 175–192.

Jobe, J.B., White, A.A., Kelley, C.L., Mingay, D.J., Sanchez, M.J., and Loftus, E.F. (1990). Recall Strategies and Memory for Health-Care Visits. The Millbank Quarterly, 68, 171–189.

Joseph, J.G., Emmons, C.-A., Kessler, R.C., Wortman, C.B., O'Brien, K., Hocker, W.T., and Schaefer, C. (1984). Coping with the Threat of AIDS. American Psychologist, 39, 1297–1302.

Kearney, A., Tourangeau, R., Shapiro, G., and Ernst, L. (1993). Coverage Improvement from Experimental Residence Questions. Proceedings of the Survey Research Methods Section, the American Statistical Association.

Krosnick, J.A. (1991). Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys. Applied Cognitive Psychology, 5, 213–236.

Krosnick, J.A. and Alwin, D.F. (1987). An Evaluation of a Cognitive Theory of Response-Order Effects. Public Opinion Quarterly, 51, 201–219.

Krosnick, J.A., Li, F., and Lehman, D.R. (1990). Conversational Conventions, Order of Information Acquisition, and the Effect of Base Rates and Individuating Information on Social Judgments. Journal of Personality and Social Psychology, 59, 1140–1152.

Krosnick, J.A. and Schuman, H. (1988). Attitude Intensity, Importance, and Certainty, and Susceptibility to Response Effects. Journal of Personality and Social Psychology, 54, 940–952.

Lessler, J., Tourangeau, R., and Salter, W. (1989). Questionnaire Design in the Cognitive Research Laboratory: Results of an Experimental Prototype. National Center for Health Statistics, Vital and Health Statistics, Series 6, No. 1, Washington, D.C.: Government Printing Office.

Loftus, E.F., Smith, K.D., Klinger, M.R., and Fiedler, J. (1992). Memory and Mismemory for Health Events. In Questions about Questions: Inquiries into the Cognitive Bases of Surveys, ed. J. Tanur. New York: Russell Sage Foundation.

McClendon, M. (1991). Acquiescence and Recency Response-Order Effects in Interview Surveys. Sociological Methods and Research, 20, 60–103.

McClendon, M.J. and Alwin, D. (1993). No-Opinion Filters and Attitude Measurement Reliability. Sociological Methods and Research, 21, 438–464.

Marquis, K.H. (1971). Effects of Race, Residence, and Selection of Respondent on the Conduct of the Interview. Chapter 12 in Working Papers on Survey Research in Poverty Areas by John B. Lansing, Stephen B. Withey, and Arthur C. Wolfe. Ann Arbor, MI: Institute for Social Research, The University of Michigan.

Martin, E.A. (1993). Response Errors in Survey Measurements of Facts. Paper presented at the 49th meeting of the International Statistical Institute, Florence, Italy, August.

Martin, E.A., Campanelli, P.C., and Fay, R.E. (1991). An Application of Rasch

Analysis to Questionnaire Design: Using Vignettes to Study the Meaning of 'Work' in the Current Population Survey. The Statistician, 40, 265–276.

Martin, E., DeMaio, T., and Campanelli, P. (1990). Context Effects for Census Measures of Race and Hispanic Origin. Public Opinion Quarterly, 54, 551–566.

Martin E.A. and Griffin, D.H. (1994). The Role of Questionnaire Design in Reducing Census Coverage Error. Paper presented at the annual meetings of the American Statistical Association, Survey Methods Research Section, Toronto, July.

Mathiowetz, N.A., and Duncan, G.J. (1988). Out of Work, Out of Mind: Response Errors in Retrospective Reports of Unemployment. Journal of Business and Economic Statistics, 6, 221–229.

Means, B. and Loftus, E.F. (1991). When Personal History Repeats Itself: Decomposing Memories for Recurring Events. Applied Cognitive Psychology, 5, 297–318.

Means, B., Nigam, A., Zarrow, M., Loftus, E.F., Donaldson, M.S. (1989). Autobiographical Memory for Health-Related Events. National Center for Health Statistics, Vital and Health Statistics, Series 6, No. 2, Washington, D.C.: Government Printing Office.

Means, B., Swan, G.E., Jobe, J.B., and Esposito, J.L. (1992). An Alternative Approach to Obtaining Personal History Data. In Measurement Errors in Surveys, eds. P. Biemer, R. Groves, L. Lyberg, N. Mathiowetz, and S. Sudman. New York: John Wiley.

Menon, G. (1993). The Effects of Accessibility of Information in Memory on Judgments of Behavioral Frequencies. Journal of Consumer Research, 20, 431–440.

Mullin, P.A., Conrad, F.G., Sander, J.E., and Herrmann, D. (1993). Cognitive Process Models of the Survey Interview. Unpublished manuscript. Bureau of Labor Statistics.

Mishler, E.G. (1986). Research Interviewing: Context and Narrative. Cambridge: Harvard University Press.

O'Reilly, J.M., Hubbard, M.L., Lessler, J.T., Biemer, P.P., and Turner, C.F. (1994). Audio and Video Computer Assisted Self-Interviewing: Preliminary Tests of New Technologies for Data Collection. Journal of Official Statistics, 10, 197–214.

Oskenberg, L., Cannell, C.F., and Kalton, G. (1991). New Strategies for Pretesting Survey Questions. Journal of Official Statistics, 7, 349–365.

Polivka, A.E. and Rothgeb, J.M. (1993). Redesigning the CPS Questionnaire. Monthly Labor Review 116, 10–28.

Presser, S. and Blair, J. (1994). Survey Pretesting: Do Different Methods Produce Different Results? In Sociological Methodology 1994, ed., P.V. Marsden. Oxford: Basil Blackwell.

Rodgers, W.L., Brown, C., and Duncan, G.J. (1993). Errors in Survey Reports of Earnings, Hours Worked, and Hourly Wages. Journal of the American Statistical Association, 88, 1208–1219.

Sanchez, M.E. (1992). Effect of Questionnaire Design on the Quality of Survey Data. Public Opinion Quarterly, 56, 206–217.

Saris, W.E. (1991). Computer-Assisted Interviewing. Newbury Park: Sage.

Schaeffer, N.C. (1994). Errors of Experience: Response Errors in Reports About Child Support and Their Implications for Questionnaire Design. In

Autobiographical Memory and the Validity of Retrospective Reports, ed. by N. Schwarz and S. Sudman. New York: Springer Verlag.

Schaeffer, N.C. (1991). Conversation with a Purpose or Conversation? Interaction in the Standardized Interview. In Measurement Errors In Surveys, eds. P. Biemer, R. Groves, L. Lyberg, N. Mathiowetz, and S. Sudman. New York: John Wiley.

Schaeffer, N.C. (1992). Hardly Ever or Constantly? Group Comparisons Using Vague Quantifiers. Public Opinion Quarterly, 55, 395–423.

Schaeffer, N.C. and Maynard, D. (1995). Occasioning Intervention: Interactional Resources for Comprehension in Standardized Survey Interviewing. Paper presented at the International Conference on Survey Measurement and Process Quality, Bristol, England.

Schaeffer, N.C. and Maynard, D.W. (1995). From Paradigm to Prototype and Back Again: Interactive Aspects of 'Cognitive Processing' in Survey Interviews. In Methods for Determining Cognitive Processes in Answering Questions, edited by N. Schwarz and S. Sudman. San Francisco: Jossey Bass.

Schaeffer, N.C. and Thomson, E. (1992). The Discovery of Grounded Uncertainty: Developing Standardized Questions about Strength of Fertility Motivation. In Sociological Methodology 1992, ed., P.V. Marsden. Washington, D.C., the American Statistical Association.

Schuman, H. and Presser, S. (1981). Questions and Answers in Attitude Surveys. Orlando: Academic Press.

Schwarz, N. (1994). Judgment in Social Context: Biases, Shortcomings, and the Logic of Conversation. In Advances in Experimental Social Psychology, Volume 26, ed. M. Zanna, San Diego, CA: Academic Press.

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

Schwarz, N. and Sudman, S. eds. (1994). Autobiographical Memory and the Validity of Retrospective Reports. New York: Springer Verlag.

Schwarz, N., and Sudman, S. eds. (forthcoming). Methods for Determining Cognitive Processes in Answering Questions. San Francisco: Jossey Bass.

Smith, A.F. (1991). Cognitive Processes in Long-Term Dietary Recall. National Center for Health Statistics, Vital and Health Statistics, Series 6, No. 4, Washington, D.C.: Government Printing Office.

Strack, F., and Schwarz, N. (1992). Communicative Influences in Standardized Question Situations: The Case of Implicit Collaboration. In Language, Interaction, and Social Cognition. eds. K. Fielder and G. Semin. Beverly Hills: Sage.

Suchman, L. and Jordan, B. (1990). Interactional Troubles in Face-to-Face Survey Interviews. Journal of the American Statistical Association, 85, 232–241.

Sudman, S., Warneke, R., Johnson, T., O'Rourke, D., and Davis, A. (1994). Cognitive Aspects of Reporting Cancer Prevention Examinations and Tests. National Center for Health Statistics, Series 6, No. 7, Washington, D.C.: Government Printing Office.

Sudman, S. and Bradburn, N.M. (1974). Response Effects in Surveys. Chicago: Aldine.

Tanur, J., ed., 1992. Questions about Questions: Inquiries into the Cognitive Bases of Surveys. New York: Russell Sage Foundation.

Tourangeau, R. (1992). Context Effects on Responses to Attitude Questions: Attitudes as Memory Structures. In Context Effects in Social and Psychological Research, ed. N. Schwarz and S. Sudman. New York: Springer Verlag.

Tourangeau, R. (1984). Cognitive Sciences and Survey Method. In Cognitive Aspects of Survey Methodology: Building a Bridge Between Disciplines, eds. T.B. Jabine, M.L. Straf, J.M. Tanur, and R. Tourangeau, Washington, D.C.: National Academy Press.

Tourangeau, R., Rasinski, K.A., and Bradburn, N. (1991). Measuring Happiness in Surveys: A Test of the Subtraction Hypothesis. Public Opinion Quarterly, 55, 255–266.

Turner, C.F. and Martin, E. eds. (1984). Surveying Subjective Phenomena. New York: Russell Sage.

Watson, D. (1992). Correcting for Acquiescent Response Bias in the Absence of a Balanced Scale: An Application to Class Consciousness. Sociological Methods and Research, 21, 52–88.

Wright, D.B., Gaskell, G.E., and O'Muircheartaigh, C.A. (1994). How Much is "Quite a bit"? Mapping Between Numerical Values and Vague Quantifiers. Applied Cognitive Psychology, 8, 479–496.

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