1

Cognitive Aspects of Surveys and This Volume

JUDITH M. TANUR

Since sample survey technology originated in studies of London's poor in the late nineteenth century, it has undergone continual refinements. Many of these refinements have improved sampling and estimation procedures. In the 1930s the United States government agencies adopted probability sampling methods as a means of ensuring the representativeness of a sample. But probability sampling did not replace the convenience and quota sampling strategies of prior decades in many other survey organizations until these strategies were proved untrustworthy when the pre-election polls of 1948 "elected" Thomas Dewey but the voters chose Harry Truman (see Fienberg and Tanur, 1989). By now, although issues of improving the efficiency of sampling and the accuracy of estimation remain active areas of investigation, researchers have sufficiently perfected probability sampling methods for telephone, mail, and face-to-face surveys that the errors such methods introduce into the results of surveys are controllable and quantifiable.

The 1960s and 1970s saw the development of new analytical technologies, especially the hardware and software associated with high-speed electronic computers. These developments permitted more people to do more surveys more cheaply and analyze them more quickly than ever before possible. They also meant that more people became respondents to both legitimate and less legitimate surveys than ever before—gone forever were the days when Phyllis McGinley (1954), after complaining that no polltaker had yet sought her opinions, could beg:

> Before the unpolled generations trample me,
> Won't someone sample me?

Indeed, during the 1970s survey researchers found that sampled respondents were increasingly refusing to grant interviews. Such refusals raise the cost of

surveys because they necessitate callbacks and degrade the accuracy of results if chosen respondents cannot be persuaded to cooperate. If refusals were occurring, even in part, because of the content of survey questionnaires, it behooved researchers to improve those questionnaires to make them less onerous.

During the same period, government agencies and academic researchers initiated longitudinal surveys, and the United States government initiated large survey-based social experiments to discover the effects of such projected reforms as a negative income tax and universal health insurance. These joined such important government surveys as the Current Population Survey carried out monthly to estimate the nation's unemployment rate, the National Crime Survey collecting data on victimizations to supplement the police-report-based Uniform Crime Reports, and the National Health Interview Survey providing data on prevalence of illnesses and their effects. This proliferation of surveys has made them part of the very fabric of our lives, providing data for academic research and for crucial government policy. But questions were arising: Were respondents able to recall accurately victimizations, visits to doctors, whether they had looked for work during the previous four weeks, and other autobiographical events that survey interviewers were asking them to report? How valid were these data on which academic research and government policy were increasingly based, and what could be done to make surveys yield more valid data? A panel of the National Academy of Sciences reflected these concerns in a report on the National Crime Survey (Penick and Owens, 1976).

Although issues of rising refusal rates and costs of surveys and of the validity of data used for policy purposes made these questions especially urgent, they were not new. For many years, survey research methodologists had been concerned with such problems of questionnaire construction and administration as the effects of interviewers, wording, open versus closed questions, the existence of middle or don't-know alternatives, question ordering, and failures of recall. But by the middle of the 1970s, the social science community reacted to the urgency by increasingly turning its attention to that part of the survey enterprise that remains today an art, but one that we shall see has lately been augmented by a better understanding of the cognitive processes that underlie the process of asking and answering questions.

One response of the social science community was to hold a series of conferences. The Department of the Navy had placed a prohibition on surveys in the early 1970s, a time when tensions arising from the changing composition of naval personnel made it crucial to measure and understand attitudes of those serving in the navy. The Office of Naval Research jointly with the Navy Personnel Research and Development Center sponsored an interdisciplinary conference on Perspectives on Attitude Assessment Surveys

and Their Alternatives (Sinaiko and Broedling, 1976). In 1980 Albert Biderman organized a conference for the Bureau of Social Science Research (Biderman, 1980) that brought together statisticians, cognitive psychologists, and survey researchers to focus on the National Crime Survey, and in 1983 the Committee on National Statistics of the National Research Council organized an Advanced Research Seminar on Cognitive Aspects of Survey Methodology (and thereby coined the acronym CASM with the express notion that the purpose of the seminar was to attempt to build a bridge over deep interdisciplinary chasms). Spawning many research proposals and inspiring the establishment of such institutional arrangements to nurture interdisciplinary research as cognitive laboratories in government statistical agencies and the Social Science Research Council (SSRC) Committee on Cognition and Survey Research, the CASM conference stands as a landmark in the history of the effort to bring together the cognitive sciences and survey research. (See Jabine, Straf, Tanur, and Tourangeau, 1984.) It began an effort that aims to encourage and maintain a dialogue between survey methodologists and researchers in the cognitive sciences in hopes of throwing new light on old survey problems and perhaps ultimately resolving them, opening up totally new opportunities for verifying or modifying laboratory-based cognitive theories in more naturalistic test beds, and creating a new interdiscipline with a research agenda of its own. The Preface to this volume contains a more detailed history of this movement.

This volume is an attempt by the SSRC Committee on Cognition and Survey Research to weave together some strands of research and theory that have developed out of part of the movement that aims to understand the cognitive bases of survey responding, many nurtured in the committee's meetings, workshops, and other activities. Thus, its focus is mainly on the contributions of the cognitive sciences to understanding issues in survey research rather than on the twinned aim of the movement, the expansion of the empirical bases of the cognitive sciences. This introduction will discuss the contributions of this volume and point to some fulfilled and some as yet unfulfilled promises of the movement.

Why This Volume?

As the movement to study cognitive aspects of survey methodology enters its second decade, the time seems ripe to look backward to take stock of progress and forward to get some hint of where the movement is going. There have been other such efforts. For example, Jobe and Mingay (1991) have recently carried out a fine review of many of the empirical studies conducted under the movement's banner, Aborn (1989) has prepared an evaluation of the progress of the movement, and Hippler, Schwarz, and Sudman (1987) have

compiled a volume of work in the tradition. But most of the output of the movement to study cognitive aspects of surveys has appeared in reports, journal articles, and conference proceedings.

We see the special purpose of this book to be the tying together of theorizing and empirical research on the interface between the cognitive sciences and survey research. The organization of two of the parts makes that interplay especially salient, with an introductory chapter setting out theoretic themes and problems, followed by one or more empirical chapters that respond to aspects of the stated theme.

The roots of the movement to study cognitive aspects of surveys in a joint academic–government enterprise continue to influence the growth of efforts to join the cognitive sciences and survey research. Thus, much of the research has been carried out in government agencies or under contract from such agencies. Further, there is an emphasis, even in much research funded from sources other than government statistical agencies, on the kinds of "factual" questions usually asked in government surveys. This emphasis on factual questions pervades many chapters of this volume, but we have been careful to include a part on the measurement of attitudes as another important area where the importation of cognitive theories and methodologies has been informing the survey enterprise. Because few United States government surveys ask questions about attitudes or opinions, such research tends to be cross-disciplinary within academe rather than in government–academic partnership.

Following this Introduction, the volume contains five further parts, four dealing with cognitive domains implicated in the survey process and the last an overview of research carried out at the Bureau of Labor Statistics, presented as an example.

Part II is entitled "Meaning," reflecting the view that a respondent must share the meaning of a question intended by the survey researcher if he or she is to respond usefully. The part starts with a thematic overview by Herbert H. Clark and Michael F. Schober, "Asking Questions and Influencing Answers." The authors place the problem of meaning—issues that survey methodologists have studied under the rubrics of the effects of question wording, of response alternatives, of question ordering, and of context—in the perspective of the psycholinguistic concept of intentions. They point out how rules for making sense of ordinary conversations carry over to survey interviews, both providing the occasion for response effects and a systematic framework for studying and perhaps reducing them. We should note, however, that some of their vivid examples are more extreme demonstrations of ambiguous meaning and alternative interpretation than typically appear in surveys. An empirical study by Robert M. Groves, Nancy Fultz, and Elizabeth Martin, "Direct Questioning About Comprehension in a Survey Setting," explores how closely respondents' and researchers' intended meanings correspond.

These authors find that an open-ended question used as a follow-up gives useful information on the perceived meaning of a survey question.

Part III is entitled "Memory," reflecting the next step in a respondent's cognitive processing of a survey question: the retrieval of information. The thematic overview article by Robert W. Pearson, Michael Ross, and Robyn M. Dawes, "Personal Recall and the Limits of Retrospective Questions in Surveys" (chapter 4), gives a framework for understanding the processes that sometimes introduce errors into retrospective reports and some guidance on when such reports are more or less trustworthy. Then, three empirical papers explore ways in which problems caused by imperfect retrospective reports can be overcome, preceded by an introductory paper setting the experimental stage, "Improving Episodic Memory Performance of Survey Respondents" by Robert T. Croyle and Elizabeth F. Loftus (chapter 5). Chapters 6 and 7 explore the use of a two-time-frame procedure (pioneered by Crespi and Swinehart, 1982) in reducing telescoping, a tendency for respondents to move events forward in time, and thus report as occurring during a specified reference period events that actually took place earlier. In "Memory and Mismemory for Health Events," Elizabeth F. Loftus, Kyle D. Smith, Mark R. Klinger, and Judith Fiedler use the technique successfully, whereas in "Attempts to Improve the Accuracy of Self-Reports of Voting," Robert P. Abelson, Elizabeth F. Loftus, and Anthony G. Greenwald find no reduction of overreporting of voting when using a two-time-frame procedure. Loftus et al. also find that directed backward recall (reverse chronological order) is significantly better for retrieving visits to healthcare providers of the respondent's spouse, and slightly better for recalling the respondent's own visits. The remaining chapter in part III on memory, "Applying Cognitive Theory in Public Health Investigations: Enhancing Food Recall with the Cognitive Interview" (chapter 8), by Ronald P. Fisher and Kathryn L. Quigley, explores another technique that has proved useful in the laboratory for overcoming problems of retrospection and suggests ways it might be useful in field applications in surveys.

Part IV deals with the reporting of what is retrieved, opening up issues of social desirability, attitude strength, and unobtrusive measures. We have chosen to deal with these problems of expression only in the context of questions about attitudes. Robert P. Abelson presents a thematic overview on "Opportunities in Survey Measurement of Attitudes" (chapter 9), and two empirical chapters follow. The issue of the correspondence between attitudes and behavior is one of long standing (see, for example, Deutscher, 1973, and the discussion in Schuman and Presser, 1981). These two chapters treat the issue systematically from a cognitive point of view. Jon A. Krosnick and Robert P. Abelson make "The Case for Measuring Attitude Strength in Surveys" (chapter 10), presenting evidence that strong attitudes are more likely than weak ones to covary with other variables, including behavior, and

thus they urge that the extra questions (and hence interview time) needed to measure attitude strength be routinely invested in surveys. John F. Dovidio and Russell H. Fazio present in chapter 11 "New Technologies for the Direct and Indirect Assessment of Attitudes." They approach the problem of the correspondence between attitudes and behavior systematically, arguing that the more likely an individual's attitude is to be activated from memory when he or she encounters an attitude object, the more likely he or she is to act in accordance with that attitude, at least when social desirability is not involved. Further, they argue that one can measure the accessibility of an attitude by the latency of its self-report; the more rapid the response, the more accessible the attitude.

Part V on "Social Interaction" is comprised of a single chapter, "Validity and the Collaborative Construction of Meaning in Face-to-Face Surveys" (chapter 12) by Lucy Suchman and Brigitte Jordan. An earlier version of this chapter appeared in the *Journal of the American Statistical Association* (from which it is reprinted with permission). Suchman and Jordan raise questions about the rigid standardization imposed on the survey interview in the pursuit of reliable data. They suggest that because such standardization violates the usual norms of conversational behavior, it may well endanger the validity of data thus collected. The *Journal of the American Statistical Association* publication included several commentaries on the Suchman and Jordan work. Although we republish neither those commentaries nor the response by Suchman and Jordan, the version of the work appearing here takes many of those commentaries into account.

Part VI, our final section, "Government Applications," presents an overview in chapter 13 by Cathryn S. Dippo and Janet L. Norwood of empirical work recently carried out in the Collection Procedures Research Laboratory at the Bureau of Labor Statistics (BLS). Three United States government statistical agencies now operate cognitive laboratories. The committee wanted to present the flavor of the research going on in those laboratories, but for reasons of space in this volume was reluctant to try to survey them all. BLS was chosen as representative for the simplest of reasons: Janet Norwood, Commissioner of Labor Statistics, was a member of our committee and was familiar with the special concerns motivating this volume.

Promises Fulfilled and Pending[1]

We believe that there is general agreement that the establishment of the cognitive laboratories in government agencies is a major contribution of the

1. Much of the evaluative material in this Introduction has been excerpted and adapted from Tanur and Fienberg (1990).

movement to bring together the cognitive sciences and survey research. The importation of tools from the cognitive sciences and the development of a new level of awareness of the implications of the ideas that emanate from the use of these tools have enriched United States government survey enterprises. These enterprises have a long history of methodological care and experimentation; but traditionally methodological survey experiments were carried out primarily in the form of full-scale field tests. The cognitive laboratories in the government agencies now use such tools as think-aloud protocols and cognitive interviewing with small numbers of subjects to do early pretesting and to secure insight into redesign options, sometimes even options favored by previous field testing (see, for example, Tucker, Miller, Vitrano, and Doddy, 1989). Of course, field tests of innovations are crucial before a change is made in an operational survey, and an experiment properly embedded in a survey is the way such a field test should be carried out (see Fienberg and Tanur, 1988, 1989), but this new approach of going back and forth between the laboratory and the field surely adds flexibility and perhaps reduces costs.

The CASM Seminar envisaged a two-way street. Not only would the insights of the cognitive sciences shed new light on problems of survey research and perhaps lead to their eventual solution, but also survey research would open a wider laboratory for testing theories of the cognitive sciences. Jobe and Mingay (1991) and Aborn (1989) point to ways in which the movement has benefited the cognitive sciences, and cognitive psychology in particular. They note, among other things, the unanticipated ability of cognitive psychologists to take theories generated in the analysis of surveys and survey-based experiments into the laboratory for testing as well as their more anticipated ability to test laboratory-generated theories in the field. So far the results seem to be more in the direction of disproof rather than confirmation. For example, Bradburn, Rips, and Shevell (1987) hypothesized that telescoping is the result of clarity of memory (arising from the vividness or salience of an event) misleading a respondent who is using an availability heuristic to gauge the recency of an event. This hypothesis was not supported in a laboratory experiment by Thompson, Skowronski, and Lee (1988), in which misdating was not related to memorableness as rated by subjects at the time of recall. This refutation points to a system of mechanisms more complicated than those previously envisaged. On the other hand, the laboratory studies of Loftus and Fathi (1985) and those reported in this volume by Loftus, Smith, Klinger, and Fiedler suggest that recall is more efficient in a backward direction (most recent event first). Additional experiments that include the alternative of free recall, however, and the work of Jobe et al. (1991) on the National Health Interview Survey/National Medical Expenditures Survey Linkage Field Test, found that free recall of doctors' visits was at least as accurate as recall in which the respondents were

instructed as to order. Similarly, in this volume Loftus et al. found that the two-time-frame procedure was useful in aiding recall of healthcare visits; but Abelson, Loftus, and Greenwald had no success with the technique in curbing overreporting of voting. While these examples suggest that cognitively based research has just begun to yield cumulation by confirmation, there does seem to be more continuity in the research than before the CASM movement existed, both to provide a structure and to bring together researchers who might otherwise not have talked to one another.

We see systematization as another contribution of the CASM movement. As old problems that have plagued the field of survey research are explained to and explored by investigators with training in the cognitive sciences, the new perspectives these recruits from other disciplines bring to bear inspire new models and suggest new avenues of empirical research. For example, in this volume we have seen that Clark and Schober look at response effects in the psycholinguistic context of speakers' intentions; Pearson, Ross, and Dawes offer a framework for understanding when and why retrospective reports are likely to be valid; and Suchman and Jordan help us to understand systematically the kinds of problems that standardization may be introducing into survey data collection. The research of Dovidio and Fazio draws upon previous psychological work on the distinction between spontaneous and deliberate behaviors and uses these ideas to shape our understanding of the link between the expression of attitudes and subsequent behavior, mediated by the social desirability of the attitudes. Whereas there is a substantial survey research literature that looks at the link between attitudes and behavior, and part of it invokes the concept of social desirability, that previous literature does not give an integrated perspective rooted in the accessibility of memory. Indeed, at the joint Office of Naval Research/Navy Personnel Research and Development Center Conference in 1975, James Dabbs suggested some physiological measures of attitudes and said explicitly, "I have no rigorous derivation linking arousal to interest, affect, and stress . . ." (1976, p. 159).

Surely, this systematizing effect is not peculiar to the work presented in this volume. For example, Tanur and Fienberg (1990) cite at length the work of Huttenlocher, Hedges, and Bradburn (1990), which explores the issue of telescoping that has long concerned survey researchers. Huttenlocher et al. draw on insights from the psychological literature to construct a model for reporting errors that takes into account effects due to bounding as well as effects associated with rounding to various types of culturally prototypic values for the number of days elapsed time since an autobiographical event. The heaping-up of reported events at boundaries of reference periods as well as at selected special values is not a new observation, but the new perspective that Huttenlocher et al. bring offers an integrated interpretation of the previously observed phenomena and a framework in which further discussions about them can take place.

Thus, we believe that the movement to bring together the cognitive sciences and survey research has so far raised more questions than it has answered. Several of the empirical findings remain contradictory. But in these contradictions lie the seeds for a better understanding of cognitive processes and improved means of asking questions in surveys. Insights and tools from the cognitive sciences have established a firm role in guiding research on survey procedures. We have not yet, however, been able to use cognitive theories directly to shape survey innovations or to predict the results of using innovative variations on established procedures. We must hope that such theoretical advances lie in the future. Surely, if survey researchers expected that they would find a panacea in the cognitive sciences that would quickly solve all the problems they had been wrestling with for decades, they have been sorely disappointed in the results of the movement so far. But if their hopes were more modest, if they had considered the development of new perspectives and systematic ways of looking at measurement errors in surveys to be progress in the improvement of the survey enterprise, and if they envisage eventual but not immediate cognitive theories of surveys, then the verdict is different. In that case, we believe, the movement to explore cognitive aspects of surveys has offered, and continues to offer, promise.

References

ABORN, M. (1989) Is CASM bridging the chasm? Evaluation of an experiment in cross-disciplinary survey research. Paper presented at the American Statistical Association 1989 Winter Conference, San Diego, CA, January 4–6.

BIDERMAN, A. (1980) *Report of a Workshop on Applying Cognitive Psychology to Recall Problems of the National Crime Survey*. Washington, DC: Bureau of Social Science Research.

BRADBURN, N. M., RIPS, L. J., and SHEVELL, S. K. (1987) Answering autobiographical questions: The impact of memory and inference on surveys. *Science* 236, 157–161.

CRESPI, I., and SWINEHART, J. W. (1982) Some effects of sequenced questions using different time intervals on behavioral self-reports: A field experiment. Paper presented at the Annual Conference of the American Association for Public Opinion Research, May.

DABBS, J. M., JR. (1976) Physiological and physical measures of attitudes. In SINAIKO, H. W., and BROEDLING, L. A. (eds.). *Perspectives on Attitude Assessment Surveys and Their Alternatives*. Champaign, IL: Pendleton, pp. 159–170.

DEUTSCHER, I. (1973) *What We Say/What We Do*. Glenview, IL: Scott, Foresman.

FIENBERG, S. E., and TANUR, J. M. (1988) From the inside out and the outside in: Combining experimental and sampling structures. *Canadian Journal of Statistics* 19, 135–151.

——— (1989) Combining cognitive and statistical approaches to survey design. *Science* 243, 1017–1022.

HIPPLER, H. J., SCHWARZ, N., and SUDMAN, S. (1987) *Social Information Processing and Survey Methodology*. New York: Springer-Verlag.

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, DC: National Academy Press.

JOBE, J. B., and MINGAY, D. J. (1991) Cognition and survey measurement: History and overview. *Applied Cognitive Psychology* 5, 175–193.

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. *Millbank Memorial Fund Quarterly/Health and Society* 68, 171–189.

LOFTUS, E. F., and FATHI, D. (1985) Retrieving multiple autobiographical memories. *Social Cognition* 3, 280–295.

MCGINLEY, P. (1954) "The Forgotten Woman." In *The Love Letters of Phyllis McGinley*. New York: Viking Press, pp. 87–89.

PENICK, B. K., and OWENS, M. E. B. (1976) *Surveying Crime*. Washington, DC: National Academy of Sciences.

SCHUMAN, H., and PRESSER, S. (1981) *Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording, and Context*. New York: Academic Press.

SINAIKO, H. W., and BROEDLING, L. A., eds. (1976) *Perspectives on Attitude Assessment Surveys and Their Alternatives*. Champaign, IL: Pendleton.

TANUR, J. M., and FIENBERG, S. E. (1990) Cognitive aspects of surveys: Yesterday, today, and tomorrow. Paper presented at the International Conference on Measurement Errors in Surveys, Tucson, AZ, November 13.

THOMPSON, C. P., SKOWRONSKI, J. J., and LEE, D. J. (1988) Telescoping in dating naturally occurring events. *Memory and Cognition* 16, 461–468.

TUCKER, C., MILLER, L., VITRANO, F., and DODDY, J. (1989) Cognitive issues and research on the Consumer Expenditure Diary Survey. Paper presented at the annual conference of the American Association for Public Opinion Research, May.