

SURVEY RESEARCH

Jon A. Krosnick

Department of Psychology, Ohio State University, Columbus, Ohio 43210;
e-mail: krosnick@osu.edu

KEY WORDS: surveys, interviewing, polls, questionnaires, pretesting

ABSTRACT

For the first time in decades, conventional wisdom about survey methodology is being challenged on many fronts. The insights gained can not only help psychologists do their research better but also provide useful insights into the basics of social interaction and cognition. This chapter reviews some of the many recent advances in the literature, including the following: New findings challenge a long-standing prejudice against studies with low response rates; innovative techniques for pretesting questionnaires offer opportunities for improving measurement validity; surprising effects of the verbal labels put on rating scale points have been identified, suggesting optimal approaches to scale labeling; respondents interpret questions on the basis of the norms of everyday conversation, so violations of those conventions introduce error; some measurement error thought to have been attributable to social desirability response bias now appears to be due to other factors instead, thus encouraging different approaches to fixing such problems; and a new theory of satisficing in questionnaire responding offers parsimonious explanations for a range of response patterns long recognized by psychologists and survey researchers but previously not well understood.

CONTENTS

INTRODUCTION	538
SAMPLING AND RESPONSE RATES	539
PRETESTING	541
RIGID INTERVIEWING VERSUS CONVERSATIONAL INTERVIEWING	542
QUESTIONNAIRE DESIGN	543
<i>Open versus Closed Questions</i>	543
<i>Labeling of Rating-Scale Points</i>	544
<i>Role of Conversational Conventions</i>	545

<i>Social Desirability Bias</i>	545
<i>Optimizing versus Satisficing</i>	546
SUMMARY	559
CONCLUSION	560

INTRODUCTION

These are exciting times for survey research. The literature is bursting with new insights that demand dramatic revisions in the conventional wisdom that has guided this research method for decades. Such dramatic revisions are nothing new for survey researchers, who are quite experienced with being startled by an unexpected turn of events that required changing their standard practice. Perhaps the best known such instance involved surveys predicting US election outcomes, which had done reasonably well at the start of the twentieth century (Robinson 1932). But in 1948 the polls predicted a Dewey victory in the race for the American presidency, whereas Truman actually won easily (Mosteller et al 1949). At fault were the nonsystematic methods used to generate samples of respondents, so we learned that representative sampling methods are essential to permit confident generalization of results.

Such sampling methods soon came into widespread use, and survey researchers settled into a “standard practice” that has stood relatively unchallenged until recently (for lengthy discussions of the method, see Babbie 1990; Lavrakas 1993; Weisberg et al 1996). This standard practice included not only the notion that systematic, representative sampling methods must be used, but also that high response rates must be obtained and statistical weighting procedures must be imposed to maximize representativeness. Furthermore, although face-to-face interviewing was thought to be the optimal method, the practicalities of telephone interviewing made it the dominant mode since the mid-1980s. Self-administered mail surveys were clearly undesirable, because they typically obtained low response rates. And although a few general rules guided questionnaire design (e.g. Parten 1950), most researchers viewed it as more of an art than a science. There is no best way to design a question, said proponents of this view; although different phrasings or formats might yield different results, all are equally informative in providing insights into the minds of respondents.

Today, this conventional wisdom is facing challenges from many directions. We have a refreshing opportunity to rethink how best to implement surveys and enhance the value of research findings generated using this method. This movement has three valuable implications for psychology. First, researchers who use the survey method to study psychological phenomena stand to benefit, because they can enhance the validity of their substantive results by using new methodologies, informed by recent lessons learned. Second, these

insights provide opportunities to reconsider past studies, possibly leading to recognize that some apparent findings were illusions. Third, many recent lessons provide insights into the workings of the human mind and the unfolding of social interaction. Thus, these insights contribute directly to the building of basic psychological theory.

Because recent insights are so voluminous, this chapter can describe only a few, leaving many important ones to be described in future *Annual Review of Psychology* chapters. One significant innovation has been the incorporation of experiments within surveys, thus permitting strong causal inference with data from representative samples. Readers may learn about this development from a chapter in the *Annual Review of Sociology* (Sniderman & Grob 1996). The other revelations, insights, and innovations discussed here are interesting because they involve the overturning of long-standing ideas or the resolution of mysteries that have stumped researchers for decades. They involve sampling and response rates, questionnaire pretesting, interviewing, and questionnaire design.

SAMPLING AND RESPONSE RATES

One hallmark of survey research is a concern with representative sampling. Scholars have, for many years, explored various methods for generating samples representative of populations, and the family of techniques referred to as probability sampling methods do so quite well (e.g. Henry 1990, Kish 1965). Many notable inaccuracies of survey findings were attributable to the failure to employ such techniques (e.g. Laumann et al 1994, Mosteller et al 1949). Consequently, the survey research community believes that representative sampling is essential to permit generalization from a sample to a population.

Survey researchers have also believed that, for a sample to be representative, the survey's response rate must be high. However, most telephone surveys have difficulty achieving response rates higher than 60%, and most face-to-face surveys have difficulty achieving response rates higher than 70% (Brehm 1993). Response rates for most major American national surveys have been falling during the last four decades (Brehm 1993, Steeh 1981), so surveys often stop short of the goal of a perfect response rate.

In even the best academic surveys, there are significant biases in the demographic and attitudinal composition of samples obtained. Brehm (1993) showed that, in the two leading, academic national public-opinion surveys (the National Election Studies and the General Social Surveys), certain demographic groups have been routinely represented in misleading numbers. Young and old adults, males, and people with the highest income levels are underrepresented, whereas people with the lowest education levels are overrepresented. Likewise, Smith (1983) found that people who do not participate

in surveys are likely to live in big cities and work long hours. And Cialdini et al (unpublished manuscript) found that people who agreed to be interviewed were likely to believe it is their social responsibility to participate in surveys, to believe that they could influence government and the world around them, and to be happy with their lives. They were also unlikely to have been contacted frequently to participate in surveys, to feel resentful about being asked a personal question by a stranger, and to feel that the next survey in which they will be asked to participate will be a disguised sales pitch. According to conventional wisdom, the higher the response rate, the less these and other sorts of biases should be manifest in the obtained data.

In the extreme, a sample will be nearly perfectly representative of a population if a probability sampling method is used and if the response rate is 100%. But it is not necessarily true that representativeness increases monotonically with increasing response rate. Remarkably, recent research has shown that surveys with very low response rates can be more accurate than surveys with much higher response rates. For example, Visser et al (1996) compared the accuracy of self-administered mail surveys and telephone surveys forecasting the outcomes of Ohio statewide elections over a 15-year period. Although the mail surveys had response rates of about 20% and the telephone surveys had response rates of about 60%, the mail surveys predicted election outcomes much more accurately (average error = 1.6%) than did the telephone surveys (average error = 5.2%). The mail surveys also documented voter demographic characteristics more accurately. Therefore, having a low response rate does not necessarily mean that a survey suffers from a large amount of nonresponse error.

Greenwald et al (AG Greenwald, unpublished manuscript) suggested one possible explanation for this finding. They conducted telephone surveys of general public samples just before elections and later checked official records to determine whether each respondent voted. The more difficult it was to contact a person to be interviewed, the less likely he or she was to have voted. Therefore, the more researchers work at boosting the response rate, the less representative the sample becomes. Thus, telephone surveys would forecast election outcomes more accurately by accepting lower response rates, rather than aggressively pursuing high response rates.

Studies of phenomena other than voting have shown that achieving higher response rates or correcting for sample composition bias do not necessarily translate into more accurate results. In an extensive set of analyses, Brehm (1993) found that statistically correcting for demographic biases in sample composition had little impact on the substantive implications of correlational analyses. Furthermore, the substantive conclusions of a study have often remained unaltered by an improved response rate (e.g. Pew Research Center 1998, Traugott et al 1987). When substantive findings did change, no evidence

allowed researchers to assess whether findings were more accurate with the higher response rate or the lower one (e.g. Traugott et al 1987). In light of Visser et al's (1996) evidence, we should not presume the latter findings were less valid than the former.

Clearly, the prevailing wisdom that high response rates are necessary for sample representativeness is being challenged. It is important to recognize the inherent limitations of nonprobability sampling methods and to draw conclusions about populations or differences between populations tentatively when nonprobability sampling methods are used. But when probability sampling methods are used, it is no longer sensible to presume that lower response rates necessarily signal lower representativeness.

PRETESTING

Questionnaire pretesting identifies questions that respondents have difficulty understanding or interpret differently than the researcher intended. Until recently, conventional pretesting procedures were relatively simplistic. Interviewers conducted a small number of interviews (usually 15–25), then discussed their experiences in a debriefing session (e.g. Bischooping 1989, Nelson 1985). They described problems they encountered (e.g. identifying questions requiring further explanation or wording that was confusing or difficult to read) and their impressions of the respondents' experiences in answering the questions. Researchers also looked for questions that many people declined to answer, which might suggest the questions were badly written. Researchers then modified the survey instrument to increase the likelihood that the meaning of each item was clear and that the interviews proceeded smoothly.

Conventional pretesting clearly has limitations. What constitutes a "problem" in the survey interview is often defined rather loosely, so there is potential for considerable variance across interviewers in terms of what is reported during debriefing sessions. Debriefings are relatively unstructured, which might further contribute to variance in interviewers' reports. And most important, researchers want to know about what went on in respondents' minds when answering questions, and interviewers are not well positioned to characterize such processes.

Recent years have seen a surge of interest in alternative pretesting methods, one of which is *behavior coding* (Cannell et al 1981, Fowler & Cannell 1996), in which an observer monitors pretest interviews (either live or taped) and notes events that occur during interactions between interviewers and respondents that constitute deviations from the script (e.g. the interviewer misreads the questionnaire, or the respondent asks for more information or provides an unclear or incomplete initial response). Questions that elicit frequent deviations are presumed to require modification.

Another new method is *cognitive pretesting*, which involves asking respondents to “think aloud” while answering questions, verbalizing whatever comes to mind as they formulate responses (e.g. Bickart & Felcher 1996, DeMaio & Rothgeb 1996, Forsyth & Lessler 1991). This procedure is designed to assess the cognitive processes by which respondents answer questions, thus providing insight into the way each item is comprehended and the strategies used to devise answers. Respondent confusion and misunderstandings can readily be identified in this way.

These three pretesting methods focus on different aspects of the survey data collection process and differ in terms of the kinds of problems they detect, as well as in the reliability with which they detect these problems. Presser & Blair (1994) demonstrated that behavior coding is quite consistent in detecting apparent respondent difficulties and interviewer problems. Conventional pretesting also detects both sorts of problems, but less reliably. In fact, the correlation between the apparent problems diagnosed in independent conventional pretesting trials of the same questionnaire can be remarkably low. Cognitive interviews also tend to exhibit low reliability across trials and to detect respondent difficulties almost exclusively. But low reliability might reflect the capacity of a particular method to continue to reveal additional, equally valid problems across pretesting iterations, a point that future research must address.

RIGID INTERVIEWING VERSUS CONVERSATIONAL INTERVIEWING

One prevailing principle of the survey method is that the same questionnaire should be administered identically to all respondents (e.g. Fowler & Mangione 1990). If questions are worded or delivered differently to different people, then researchers cannot be certain about whether differences between the answers are due to real differences between the respondents or are due to the differential measurement techniques employed. Since the beginning of survey research this century, interviewers have been expected to read questions exactly as researchers wrote them, identically for all respondents. If respondents expressed uncertainty and asked for help, interviewers avoided interference by saying something like “it means whatever it means to you.”

Some critics have charged that this approach compromises data quality instead of enhancing it (Briggs 1986, Mishler 1986, Suchman & Jordan 1990, 1992). In particular, they have argued that the meanings of many questions are inherently ambiguous and are negotiated in everyday conversation through back-and-forth exchanges between questioners and answerers. To prohibit such exchanges is to straight-jacket them, preventing precisely what is needed to maximize response validity. Schober & Conrad (1997) recently reported the first convincing data on this point, demonstrating that when interviewers were

free to clarify the meanings of questions and response choices, the validity of reports increased substantially.

This finding has important implications for technological innovations in questionnaire administration. Whereas survey questionnaires were traditionally printed on paper, most large-scale survey organizations have been using computer-assisted telephone interviewing (CATI) for the last decade. Interviewers read questions displayed on a computer screen; responses are entered immediately into the computer; and the computer determines the sequence of questions to be asked. This system can reduce some types of interviewer error and permits researchers to vary the specific questions each participant is asked on the basis of previous responses.

All this has taken another step forward recently: Interviewers conducting surveys in people's homes are equipped with laptop computers (for computer-assisted personal interviewing, or CAPI), and the entire data collection process is regulated by computer programs. In audio computer-assisted self-administered interviewing (audio CASAI), a computer reads questions aloud to respondents who listen on headphones and type their answers on computer keyboards. Thus, computers have replaced interviewers. Although these innovations have clear advantages for improving the quality and efficiency of questionnaire administration, this last shift may be problematic in light of Schober & Conrad's (1997) evidence that conversational interviewing can significantly improve data quality. Perhaps technological innovation has gone one step too far, because without a live interviewer, conversational questioning is impossible.

QUESTIONNAIRE DESIGN

Open versus Closed Questions

During the 1940s, a major dispute erupted between two survey research divisions of the US Bureau of Intelligence, the Division of Polls and the Division of Program Surveys. The former was firmly committed to asking closed-ended questions, which required people to choose among a set of provided response alternatives. The latter believed in the use of open-ended questions, which respondents answered in their own words (see Converse 1987). Paul Lazarsfeld mediated the dispute and concluded that the quality of data collected by each method seemed equivalent, so the greater cost of administering open-ended questions did not seem worthwhile (see Converse 1987). Over time, closed-ended questions have become increasingly popular, whereas open-ended questions have been asked less frequently (Smith 1987).

Recent research has shown that there are distinct disadvantages to closed-ended questions, and open-ended questions are not as problematic as they

seemed. For example, respondents tend to confine their answers to the choices offered, even if the researcher does not wish them to do so (Bishop et al 1988, Presser 1990). That is, people generally ignore the opportunity to volunteer a response and simply select among those listed, even if the best answer is not included. Therefore, a closed-ended question can only be used effectively if its answer choices are comprehensive, and this is difficult to assure.

Some people feared that open-ended questions would not work well for respondents who are not especially articulate, because they might have difficulty explaining their feelings. However, this seems not to be a problem (Geer 1988). Some people feared that respondents would be likely to answer open-ended questions by mentioning the most salient possible responses, not those that are truly most appropriate. But this, too, turns out not to be the case (Schuman et al 1986). Finally, a number of recently rediscovered studies found that the reliability and validity of open-ended questions exceeded that of closed-ended questions (e.g. Hurd 1932, Remmers et al 1923). Thus, open-ended questions seem to be more viable research tools than had seemed to be the case.

Labeling of Rating-Scale Points

Questionnaires have routinely offered rating scales with only the endpoints labeled with words and the points in between either represented graphically or labeled with numbers and not words. However, reliability and validity can be significantly improved if all points on the scale are labeled with words, because they clarify the meanings of the scale points (Krosnick & Berent 1993, Peters & McCormick 1966). Respondents report being more satisfied when more rating-scale points are verbally labeled (e.g. Dickinson & Zellinger 1980), and validity is maximized when the verbal labels have meanings that divide the continuum into approximately equal-sized perceived units (e.g. Klockars & Yamagishi 1988). On some rating dimensions, respondents presume that a “normal” or “typical” person falls in the middle of the scale, and some people are biased toward placing themselves near that point, regardless of the labels used to define it (Schwarz et al 1985).

Another recent surprise is that the numbers used by researchers to label rating-scale points can have unanticipated effects. Although such numbers are usually selected arbitrarily (e.g. an 11-point scale is labeled from 0 to 10, rather than from -5 to +5), respondents sometimes presume that these numbers were selected to communicate intended meanings of the scale points (e.g. a unipolar rating for the 0 to 10 scale and a bipolar rating for the -5 to +5 scale; Schwarz et al 1991). Consequently, a change in the numbering scheme can produce a systematic shift in responses. This suggests either that rating-scale points should be labeled only with words or that numbers should reinforce the meanings of the words, rather than communicate conflicting meanings.

Conversational Conventions

Survey researchers have come to recognize that respondents infer the meanings of questions and response choices partly from norms and expectations concerning how everyday conversations are normally conducted (Schwarz 1996). Speakers conform to a set of conventions regarding what to say, how to say it, and what not to say; these conventions make conversation efficient by allowing speakers to convey unspoken ideas underlying their utterances (e.g. Clark 1996, Grice 1975). Furthermore, listeners presume that speakers are conforming to these norms when interpreting utterances. Respondents bring these same conventions to bear when they interpret survey questions, as well as when they formulate answers (see Schwarz 1996).

Krosnick et al (1990) showed that the order in which information is provided in the stem of a question is sometimes viewed as providing information about the importance or value the researcher attaches to each piece of information. Specifically, respondents presume that researchers provide less important “background” information first and then present more significant “foreground” information later. Consequently, respondents place more weight on more recently presented information because they wish to conform to the researcher’s beliefs. From these studies and various others (see Schwarz 1996), we now know that we must guard against the possibility of unwittingly communicating information to respondents by violating conversational conventions, thus biasing answers.

Social Desirability Bias

One well-known phenomenon in survey research is overreporting of admirable attitudes and behaviors and underreporting those that are not socially respected. For example, the percentage of survey respondents who say they voted in the last election is usually greater than the percentage of the population that actually voted (Clausen 1968, Granberg & Holmberg 1991, Traugott & Katosh 1979). Furthermore, claims by significant numbers of people that they voted are not corroborated by official records. These patterns have been interpreted as evidence that respondents intentionally reported voting when they did not, because voting is more admirable than not doing so.

In fact, these two empirical patterns are not fully attributable to intentional misrepresentation. The first of the discrepancies is partly due to inappropriate calculations of population turnout rates, and the second discrepancy is partly caused by errors in assessments of the official records (Clausen 1968, Presser et al 1990). The first discrepancy also occurs partly because people who refuse to be interviewed for surveys are disproportionately unlikely to vote (Greenwald et al, unpublished manuscript) and pre-election interviews increase interest in politics and elicit commitments to vote, which become self-fulfilling

prophecies (Greenwald et al 1987, Yalch 1976). But even after controlling for all these factors, some people still claim to have voted when they did not.

Surprisingly, recent research suggests that the widely believed explanation for this fact may be wrong. Attempts to make people comfortable admitting that they did not vote have been unsuccessful in reducing overreporting (e.g. Abelson et al 1992, Presser 1990). People who typically overreport also have the characteristics of habitual voters and indeed have histories of voting in the past, even though not in the most recent election (Abelson et al 1992, Sigelman 1982, Silver et al 1986). And the accuracy of turnout reports decreases as time passes between an election and a postelection interview, suggesting that the inaccuracy occurs because memory traces of the behavior or lack thereof fade (Abelson et al 1992).

Most recently, Belli et al (unpublished manuscript) significantly reduced overreporting by explicitly alerting respondents to potential memory confusion and encouraging them to think carefully to avoid such confusion. These instructions had increasingly beneficial effects on report accuracy as more time passed between election day and an interview. This suggests that what researchers have assumed is intentional misrepresentation by respondents may be at least partly attributable instead to accidental mistakes in recall. This encourages us to pause before presuming that measurement error is due to intentional misrepresentation, even when it is easy to imagine why respondents might intentionally lie. More generally, social desirability bias in questionnaire measurement may be less prevalent than has been assumed.

Optimizing versus Satisficing

Another area of innovation involves new insights into the cognitive processes by which respondents generate answers. These insights have been publicized in a series of recent publications (e.g. Krosnick & Fabrigar 1998, Sudman et al 1996, Tourangeau et al 1998), and some of them have provided parsimonious explanations for long-standing puzzles in the questionnaire design literature. The next section reviews developments in one segment of this literature, focusing on the distinction between optimizing and satisficing.

OPTIMIZING There is wide agreement about the cognitive processes involved when respondents answer questions optimally (e.g. Cannell et al 1981, Schwarz & Strack 1985, Tourangeau & Rasinski 1988). First, respondents must interpret the question and deduce its intent. Next, they must search their memories for relevant information and then integrate that information into a single judgment (if more than one consideration is recalled). Finally, they must translate the judgment into a response by selecting one of the alternatives offered.

Each of these four steps can be quite complex, involving a great deal of cognitive work (e.g. Krosnick & Fabrigar 1998). For example, question interpretation can be decomposed into four cognitive steps, guided by a complex and extensive set of rules (e.g. Clark & Clark 1977). First, respondents bring the sounds of the words into their “working memories.” Second, they break the words down into groups, each one representing a single concept, settling on the meaning of each. If multiple interpretations exist, listeners apparently select one interpretation and proceed ahead with it, revising later only if it leads to implausible or incomprehensible conclusions. Third, respondents build the meaning of the entire question by establishing the relations among the concepts. Finally, respondents discard the original words of the question and retain their interpretations as they begin to formulate an answer.

A great deal of cognitive work is required to generate an optimal answer to even a single question, so the cumulative effort required to answer a long series of questions on a wide range of topics seems particularly substantial. A wide variety of motives may encourage expending considerable cognitive effort to do so, including desires for self-expression, interpersonal response, intellectual challenge, self-understanding, feelings of altruism, or emotional catharsis (see Warwick & Lininger 1975). Expenditure of great effort can also be motivated by desires for gratification from successful performance, to help employers improve working conditions, to help manufacturers produce better quality products, or to help governments make better informed policy decisions. To the extent that these sorts of motives inspire a person to perform the necessary cognitive tasks in a thorough and unbiased manner, a person may be said to be optimizing.

Although we hope all respondents will optimize throughout a questionnaire, this seems to be unrealistic. In fact, some people may agree to complete a questionnaire through a relatively automatic compliance process (e.g. Cialdini 1993) or because they need to fulfill a course requirement. Thus, they may agree merely to provide answers, with no intrinsic motivation toward high quality. Other respondents may satisfy their desires to provide high-quality data after answering a few questions, and become increasingly fatigued and distracted as a questionnaire progresses. Respondents then face a dilemma: They are not motivated to work hard, and the cognitive costs of hard work are burdensome. Nonetheless, the questionnaire continues to pose a seemingly unending stream of questions, suggesting that respondents are expected to expend the effort necessary to generate high-quality responses.

SATISFICING Respondents sometimes deal with this situation by shifting their response strategy (Krosnick 1991). Rather than expending the effort to generate optimal answers, respondents may compromise their standards and expend less energy. When done subtly, respondents may simply be less thor-

ough in comprehension, retrieval, judgment, and response selection. They may be less thoughtful about a question's meaning; they may search their memories less comprehensively; they may integrate retrieved information carelessly; and they may select a response imprecisely. All four steps are executed, but each one less diligently than when optimizing occurs. Instead of generating the most accurate answers, respondents settle for merely satisfactory ones. This response behavior might be termed "weak satisficing" (borrowing the term from Simon 1957).

A more dramatic approach is to skip the retrieval and judgment steps altogether. That is, respondents may interpret each question superficially and select what they believe will be a reasonable answer to the interviewer and researcher. Yet this answer is selected without referring to any internal psychological cues relevant to the attitude, belief, or event of interest. Instead, the respondent may look to the wording of the question for a cue, pointing to a response that can be easily selected and defended if necessary. If no such cue is present, the respondent may arbitrarily select an answer. This process might be termed "strong satisficing."

Respondents can use a number of possible decision heuristics to arrive at a satisfactory answer without expending substantial effort. A person might select the first reasonable response he or she encounters in a list rather than carefully processing all possible alternatives. Respondents could be inclined to accept assertions made in the questions regardless of content, rather than performing the cognitive work required to evaluate those assertions. Respondents might offer "safe" answers, such as the neutral point of a rating scale, endorsement of the status quo, or saying "don't know" so as to avoid expending the effort necessary to consider and possibly take more risky stands. In the extreme, respondents could randomly select a response from those offered by a closed-ended question.

Optimizing and strong satisficing can be thought of as anchoring the ends of a continuum indicating the degrees of thoroughness of the four response process steps. The optimizing end involves complete and effortful execution of all four steps. The strong satisficing end involves little effort in the interpretation and answer-reporting steps and no retrieval or integration at all. In between are intermediate levels of satisficing.

CONDITIONS THAT FOSTER SATISFICING The likelihood that a respondent will satisfice when answering a question may be a function of three factors (Krosnick 1991). Satisficing is more likely to occur (a) the greater the task difficulty, (b) the lower the respondent's ability, and (c) the lower the respondent's motivation to optimize. Task difficulty is a function of the difficulty of interpreting the meaning of a question and response choices, the difficulty of retrieving and manipulating information in memory, the pace at which an interviewer reads,

the occurrence of distracting events, and more. Ability is presumably greater among respondents adept at performing complex mental operations, practiced at thinking about the topic of a question, and equipped with preformulated judgments on the issue. Factors influencing a respondent's motivation to optimize include need for cognition (Cacioppo et al 1996), the personal importance of the question's topic to the respondent, beliefs about whether the questionnaire will have useful consequences, the behavior of the interviewer, and fatigue.

EXPLAINING RESPONSE ORDER EFFECTS The notion of satisficing casts new light on many past studies of questionnaire design effects, because it provides a novel and parsimonious explanation for these effects. One such effect is the order in which response alternatives are presented on people's selection among them, called response order effects. Studies have shown that presentation order does have effects, but it has not been clear when such effects occur and what their direction might be. Some studies identified primacy effects (in which response choices presented early were most likely to be selected); other studies found recency effects (in which response choices presented last were more likely to be selected), and still other studies found no order effects at all. The satisficing perspective brought order to this evidence.

To understand the satisficing explanation here, one must distinguish categorical questions from rating-scale questions. Rating-scale questions ask people to choose a descriptor from a set that represents a dimension or continuum (e.g. from "strongly agree" to "strongly disagree"). In contrast, categorical questions ask people to choose among a set that does not represent a continuum (e.g. What is the most important problem facing the country today, unemployment or inflation?).

Response order effects in categorical questions seem to be attributable to weak satisficing (see Krosnick 1991, Krosnick & Alwin 1987). When confronted with such questions, a respondent who is optimizing would carefully assess the appropriateness of each response before selecting one. In contrast, a respondent who is a weak satisficer could simply choose the first reasonable response. Exactly which alternative is most likely to be chosen depends on whether the response choices are presented visually or orally.

When choices are presented visually, either on a show card in a face-to-face interview or in a self-administered questionnaire, weak satisficing is likely to bias respondents toward selecting choices displayed early in a list. Respondents begin at the top of the list and consider each alternative individually, and their thoughts are likely to be biased in a confirmatory direction (Klayman & Ha 1987, Koriat et al 1980, Yzerbyt & Leyens 1991). Because researchers typically include response choices that are reasonable, this confirmation-biased thinking is likely to generate at least a reason or two in favor of selecting almost any alternative a respondent considers.

After considering one or two alternatives, the potential for fatigue becomes significant, as respondents' minds become cluttered with thoughts about initial alternatives. Also, fatigue may result from proactive interference, whereby thoughts about the initial alternatives interfere with and confuse thinking about later, competing alternatives (Miller & Campbell 1959). Weak satisficers can cope by thinking only superficially about later response alternatives; the confirmatory bias would thereby give the earlier items an advantage. Alternatively, weak satisficers can terminate their evaluation process altogether once they come upon a seemingly reasonable response. Again, because most answers are likely to seem reasonable, these respondents are likely to choose alternatives near the beginning of a list. Thus, weak satisficing seems likely to produce primacy effects under conditions of visual presentation.

When response alternatives are presented orally, as in face-to-face or telephone interviews, the effects of weak satisficing are more difficult to anticipate because response order effects reflect not only evaluations of each option but also the limits of memory. When alternatives are read aloud, respondents cannot process the first one extensively. Presentation of the second alternative terminates processing of the first one, usually relatively quickly. Therefore, respondents are able to devote the most processing time to the final items read; these items remain in short-term memory after interviewers pause to let respondents answer. Thus, the last options are likely to receive deeper processing dominated by generation of reasons supporting selection.

Some respondents may listen to a short list of response alternatives without evaluating any of them. Once the list is completed, they may recall the first alternative, think about it, and then progress through the list from beginning to end. Because fatigue should instigate weak satisficing relatively quickly, a primacy effect would be expected. However, because this process requires more effort than simply considering the final items in the list first, weak satisficers are unlikely to do this very often. Considering only the allocation of processing, we would anticipate both primacy and recency effects, although the latter should be more common.

These effects are likely to be reinforced by the effects of memory. Items presented early in a list are most likely to enter long-term memory (e.g. Atkinson & Shiffrin 1968), and items presented at the end are most likely to be in short-term memory immediately after the list is heard (e.g. Atkinson & Shiffrin 1968). So items presented at the beginning and end of a list are more likely to be recalled after the question is read, particularly if the list is long. Because a response alternative must be remembered to be selected, both early and late items should be more available for selection, especially among weak satisficers. Typically, short-term memory dominates long-term memory immediately after acquiring a list of information (Baddeley & Hitch 1977), so memory factors should promote recency effects more than primacy effects. Thus, in re-

sponse to orally presented questions, mostly recency effects would be expected, though some primacy effects might occur as well.

Two additional factors may govern response order effects: the plausibility of the alternatives presented and perceptual contrast effects (Schwarz & Hippler 1991, Schwarz et al 1992). If deep processing is accorded to an alternative that seems highly implausible, even respondents with a confirmatory bias in reasoning may not generate any reasons to select it. Thus, deeper processing of some alternatives may make them especially unlikely to be selected. Also, perceptual contrast may cause a moderately plausible alternative to seem less plausible if considered after a highly plausible one or more plausible if considered after a highly implausible one.

Although the results of past studies seem to offer a mishmash of results when considered as a group, systematic patterns appear when studies are separated into ones involving visual and oral presentation. Whenever a visual presentation study has uncovered a response order effect, it has always been a primacy effect (Ayidiya & McClendon 1990, Becker 1954, Bishop et al 1988, Campbell & Mohr 1950, Israel & Taylor 1990, Krosnick & Alwin 1987, Schwarz et al 1992). And in studies involving oral presentation, nearly all response order effects documented were recency effects (Berg & Rapaport 1954, Bishop 1987, Bishop et al 1988, Cronbach 1950, Krosnick 1992, Krosnick & Schuman 1988, Mathews 1927, McClendon 1986a, 1991, Schuman & Presser 1981, Schwarz et al 1992, Visser et al 1999).

If the response order effects demonstrated in these studies are caused by weak satisficing, then they should be stronger when satisficing is most likely. Indeed, these effects were stronger among respondents with relatively limited cognitive skills (Krosnick 1991; Krosnick & Alwin 1987; Krosnick et al 1996; McClendon 1986a, 1991; Narayan & Krosnick 1996). Mathews (1927) also found stronger response order effects as questions became more difficult and respondents became fatigued. Although McClendon (1986a) found no relation between the number of words in a question and the magnitude of response order effects, Payne (1949/1950) found more response order effects in questions involving more words and words that were difficult to comprehend. Also, Schwarz et al (1992) showed that a strong recency effect was eliminated when prior questions on the same topic were asked, which presumably made respondents' knowledge of the topic more accessible and thereby made optimizing easier. The only surprise was reported by Krosnick & Schuman (1988), who found that response order effects were not stronger among respondents less certain of their opinions, who considered a question's topic to be less important, or who had weaker feelings on the issue. In general, though, this evidence is consistent with the notion that response order effects are attributable to satisficing, and evidence reported by Narayan & Krosnick (1996) and Krosnick et al (1996) ties these effects to weak satisficing in particular.

Much of the logic regarding categorical questions seems applicable to ratings scales, but in a different way. Many people's dimensional attitudes and beliefs are probably not precise points, but rather are ranges or "latitudes of acceptance" (Sherif & Hovland 1961, Sherif et al 1965). If the options on a rating scale are considered sequentially, then the respondent may select the first one that falls in his or her latitude of acceptance. This would yield a primacy effect under both visual and oral presentation, because people probably quickly consider each response alternative in the order in which they are read.

Nearly all studies of response order effects in rating scales involved visual presentation, and when order effects appeared, they were nearly uniformly primacy effects (Carp 1974, Chan 1991, Holmes 1974, Johnson 1981, Payne 1971, Quinn & Belson 1969). Two oral-presentation studies of rating scales found primacy effects as well (Kalton et al 1978, Mingay & Greenwell 1989). Consistent with the satisficing notion, Mingay & Greenwell (1989) found that a primacy effect was stronger for people with more limited cognitive skills. However, they found no relation of the magnitude of the primacy effect to the speed at which interviewers read questions, despite the fact that a fast pace presumably increased task difficulty. Also, response order effects were no stronger when questions were placed later in a questionnaire (Carp 1974). Thus, the moderators of rating-scale response order effects may be different from those for categorical questions, although more research is needed to fully address this matter.

EXPLAINING ACQUIESCENCE Agree/disagree, true/false, and yes/no questions are very popular, appearing in numerous batteries developed for attitude and personality measurement (e.g. Davis & Smith 1996, Hathaway & McKinley 1940, Robinson et al 1991, Shaw & Wright 1967). They are appealing from a practical standpoint, because they are easy to write and administer. These formats are also seriously problematic, because they are susceptible to bias due to acquiescence—the tendency to endorse any assertion made in a question, regardless of its content.

Evidence of acquiescence is voluminous and consistently compelling, based on a range of different demonstration methods (for a review, see Krosnick & Fabrigar 1998). Consider agree/disagree questions. When people are given such response choices, are not asked any questions, and are told to guess what answers an experimenter is imagining, people guess "agree" much more often than "disagree." When people are asked to agree or disagree with pairs of statements stating mutually exclusive views (e.g. "I enjoy socializing" versus "I don't enjoy socializing"), answers should be strongly negatively correlated. But across more than 40 studies, the average correlation was only $-.22$. Across 10 studies, an average of 52% of people agreed with an assertion, whereas only 42% disagreed with its opposite. In another eight studies, an average of 14% more peo-

ple agreed with an assertion than expressed the same view in a corresponding forced-choice question. And averaging across seven studies, 22% agreed with both a statement and its reversal, whereas only 10% disagreed with both.

All of these methods suggest an average acquiescence effect of about 10%, and the same sort of evidence documents comparable acquiescence in true/false and yes/no questions. There is other evidence regarding these latter question formats as well (see Krosnick & Fabrigar 1998). For example, people answer yes/no and true/false factual questions correctly more often when the correct answer is yes or true. Similarly, reports of factual matters are more likely to disagree with reports of informants when the initial reports are yes answers. And when people say they are guessing at true/false questions, they say "true" more often than "false".

Among psychologists, the prevailing explanation for acquiescence is the notion that some people may be predisposed to be agreeable in all domains of social interaction, which is consistent with the literature on the "Big Five" personality traits (Costa & McCrae 1988, Goldberg 1990). Although childhood socialization experiences probably influence an adult's level of agreeableness, this trait may have genetic roots as well (Costa & McCrae 1995). And people who are high in agreeableness are presumably inclined to acquiesce in answering all questionnaires.

Sociologists have offered a different explanation, focusing on the relationship between the respondent and the interviewer, researcher, or both. When researchers and interviewers are perceived as being of higher social status, respondents may defer to them out of courtesy and respect, yielding a tendency to endorse assertions apparently made by the researchers and/or interviewers (Carr 1971, Lenski & Leggett 1960).

Acquiescence can also be explained by the notion of satisficing (Krosnick 1991). When presented with an assertion and asked to agree or disagree, some respondents may attempt to search their memories for reasons to do each. Because of the confirmatory bias in hypothesis testing, most people typically begin by seeking reasons to agree rather than disagree. If a person's cognitive skills or motivation are relatively low, he or she may become fatigued before getting to the task of generating reasons to disagree with the assertion. The person would thus be inclined to agree. This would constitute a form of weak satisficing, because respondents would compromise their effort during the retrieval and integration stages of information processing, not during question interpretation or response expression. This is consistent with the notion that people initially believe assertions, and only upon later reflection do they come to discredit some assertions that appear insufficiently justified (Clark & Chase 1972, 1974; Gilbert et al 1990).

Acquiescence might also be a result of strong satisficing. When respondents are not able or motivated to interpret questions carefully and search their

memories for relevant information, agree/disagree, true/false, and yes/no questions offer readily available opportunities for effortless selection of a plausible response. The social convention to be polite is quite powerful, and agreeing with others is more polite than disagreeing (Brown & Levinson 1987, Leech 1983). Therefore, under conditions likely to foster strong satisficing, acquiescence may occur with no evaluation of the question's assertion at all. People may simply choose to agree because it seems like the commanded and polite action to take.

These explanations of acquiescence suggest that some people should be more likely to manifest it than others, because of personalities, social status, or abilities and motivations to optimize. Indeed, some evidence suggests that individual differences in the tendency to acquiesce are quite uniform across questions and over time (see Krosnick & Fabrigar 1998). For example, the cross-sectional reliability of the tendency to agree with a large set of assertions on diverse topics is .65, averaging across dozens of studies. Over time, the tendency to acquiesce is about .75 over one month and .67 over four months. However, consistency over time is only about .35 over four years, suggesting that the relevant disposition is not as firmly fixed as some other aspects of personality.

Evidence suggesting that multiple factors cause acquiescence comes from dozens of studies correlating the tendency across different batteries of items (see Krosnick & Fabrigar 1998). Correlations between the tendency to acquiesce on different sets of items measuring different constructs on the same occasion average .34 for agree/disagree questions, .16 for yes/no questions, and .37 for true/false questions. Correlations between acquiescence on agree/disagree batteries and yes/no batteries average .24, between acquiescence on agree/disagree and true/false item sets average .36, and between yes/no and true/false acquiescence average .21. These numbers are consistent with the conclusions that (a) a general disposition to acquiesce explains only some of variance in the acquiescence a person manifests on any particular set of items, and (b) yes/no questions may manifest this tendency less than agree/disagree or true/false items. Even more striking is that acquiescence appears to result partly from a transient, moodlike state within a single questionnaire, because the closer in time two items are presented, the more likely people are to answer them with the same degree of acquiescence (Hui & Triandis 1985, Roberts et al 1976).

In line with the status differential explanation, some studies found acquiescence to be more common among respondents of lower social status (e.g. Gove & Geerken 1977, Lenski & Leggett 1960, McClendon 1991, Ross & Mirowsky 1984), but just as many other studies failed to find this relation (e.g. Calsyn et al 1992, Falzthzik & Jolson 1974, Gruber & Lehmann 1983, Ross et al 1995). In line with the personality disposition explanation, people who acqui-

esce are unusually extraverted and sociable (Bass 1956, Webster 1958), cooperative (Heaven 1983, Husek 1961), interpersonally sensitive (Mazmanian et al 1987), and tend to have an external locus of control (Mirowsky & Ross 1991); however, none of these relations is especially strong. And although some studies found that people who acquiesce in answering questionnaires were likely to conform to others' views and comply with others' requests (e.g. Bass 1958, Kuethe 1959), more studies failed to uncover these relations (e.g. Foster 1961, Foster & Grigg 1963, Small & Campbell 1960).

In contrast, a great deal of evidence is consistent with satisficing and cannot be accounted for by these other explanations. For example, acquiescence is more common among people with more limited cognitive skills (e.g. Bachman & O'Malley 1984, Clare & Gudjonsson 1993, Forehand 1962, Gudjonsson 1990, Hanley 1959, Krosnick et al 1996, Narayan & Krosnick 1996) and with less cognitive energy (Jackson 1959), and among those who do not like to think (Jackson 1959, Messick & Frederiksen 1958). Acquiescence is more common when a question is difficult to answer (Gage et al 1957, Hanley 1962, Trott & Jackson 1967), when respondents have been encouraged to guess (Cronbach 1941), after they have become fatigued (e.g. Clancy & Wachler 1971), and during telephone interviews than during face-to-face interviews (e.g. Calsyn et al 1992, Jordan et al 1980), presumably because people feel more accountable under the latter conditions. People who acquiesce are likely to manifest other forms of satisficing (discussed below), such as nondifferentiation (Goldstein & Blackman 1976, Schutz & Foster 1963) and selecting a no-opinion option (Silk 1971). Finally, studies of thought-listings and response latencies document a confirmatory bias in reasoning when people answer agree/disagree, true/false, and yes/no questions, which is at the heart of the satisficing explanation (Carpenter & Just 1975, Kunda et al 1993). The only evidence inconsistent with the satisficing perspective is that acquiescence is not more common among people for whom the topic of a question is less personally important, who have weaker feelings on the issue, or who hold their opinions with less confidence (Husek 1961, Krosnick & Schuman 1988).

EXPLAINING THE DISCREPANCY BETWEEN RATINGS AND RANKINGS The satisficing perspective proves useful in explaining the discrepancy between ratings and rankings. An important goal of survey research is to understand the choices people make between alternative courses of action or objects. One way to do so is to explicitly ask respondents to make choices by rank ordering a set of alternatives. Another approach is to ask people to rate each object individually, allowing the researcher to derive the rank order implied by the ratings. Ratings are much less time consuming than rankings (McIntyre & Ryans 1977, Reynolds & Jolly 1980, Taylor & Kinnear 1971), and people enjoy doing ratings more and are more satisfied with their validity (Elig & Frieze 1979, McIn-

tyre & Ryans 1977). Perhaps partly as a result, researchers have typically preferred to use rating questions rather than ranking questions.

However, a number of studies indicate that rankings yield higher-quality data than ratings. Respondents are more likely to make mistakes when answering rating questions, failing to answer an item more often than when ranking (Brady 1990, Neidell 1972). Rankings are more reliable (Elig & Frieze 1979, Miethe 1985, Munson & McIntyre 1979, Rankin & Grube 1980, Reynolds & Jolly 1980) and manifest higher discriminant validity than ratings (Bass & Avolio 1989, Elig & Frieze 1979, Miethe 1985, Zuckerman et al 1989). When manifesting different correlations with criterion measures, rankings evidence greater validity than ratings (Nathan & Alexander 1985, Schriesheim et al 1991, Zuckerman et al 1989).

No explanation for this discrepancy had existed before the satisficing perspective was proposed. When confronted with a battery of ratings asking that a series of objects be evaluated on a single response scale, respondents who are inclined to implement strong satisficing can simply select a reasonable point on the scale and place all the objects at that point. For example, when asked to rate the importance of a series of values (e.g. equality, freedom, and happiness) on a scale from extremely important to not at all important, a satisficing respondent can easily say they are all very important. In the satisficing rubric, this is called nondifferentiation.

Nondifferentiation is most likely to occur under the conditions thought to foster satisficing. Nondifferentiation is more common among less educated respondents (Krosnick & Alwin 1988, Krosnick et al 1996; L Rogers & AR Herzog, unpublished manuscript) and is more prevalent toward the end of a questionnaire (Coker & Knowles 1987, Herzog & Bachman 1981, Knowles 1988, Kraut et al 1975; L Rogers & AR Herzog, unpublished manuscript). Nondifferentiation is particularly pronounced among respondents low in verbal ability, for whom fatigue is presumably most taxing (Knowles et al 1989a,b). Placing rating questions later in a questionnaire makes correlations between ratings on the same scale more positive or less negative (Andrews 1984, Herzog & Bachman 1981; L Rogers & AR Herzog, unpublished manuscript), which are the expected results of nondifferentiation (see Krosnick & Alwin 1988). Not surprisingly, removing nondifferentiators makes the validity of rating data equivalent to that of ranking data (Krosnick & Alwin 1988).

EXPLAINING SELECTION OF NO-OPINION RESPONSE OPTIONS Another application of satisficing is in explaining the effect of a no-opinion (NO) option. When researchers ask questions about subjective phenomena, they usually presume that respondents' answers reflect information or opinions that they previously had stored in memory. If a person does not have a preexisting opin-

ion, a question presumably prompts him or her to draw on relevant beliefs or attitudes in order to concoct a reasonable, albeit new, belief or evaluation (e.g. Zaller & Feldman 1992). Consequently, whether based on a preexisting judgment or a newly formulated one, responses presumably reflect the individual's belief or orientation.

When people are asked about an object about which they have no knowledge, researchers hope that respondents will say that they have no opinion, are not familiar with the object, or do not know how they feel about it. But if a question's wording suggests that respondents should have opinions, they may not wish to appear uninformed and may therefore give an arbitrary answer (Converse 1964, Schwarz 1996). Indeed, respondents have been willing to offer opinions about obscure or purely fictitious objects (Bishop et al 1986, Ehrlich & Rinehart 1965, Gill 1947, Hartley 1946, Hawkins & Coney 1981, Schuman & Presser 1981). To reduce such behavior, some survey experts have recommended that NO options routinely be offered (e.g. Bogart 1972, Converse & Presser 1986, Payne 1950, Vaillancourt 1973).

Many more respondents say they have no opinion on an issue when this option is explicitly offered than when they must volunteer it on their own (Ayidiya & McClendon 1990; Bishop et al 1980; Kalton et al 1978; McClendon 1986b, 1991; McClendon & Alwin 1993; Presser 1990; Schuman & Presser 1981). And the propensity to offer opinions about obscure or fictitious objects is significantly reduced by explicitly offering a NO option (Schuman & Presser 1981).

People who select NO responses have characteristics suggesting that they are least likely to have formed real opinions. For example, such responses are offered more often by people with relatively limited cognitive skills (Bishop et al 1980, Gergen & Back 1965, Narayan & Krosnick 1996, Sigelman 1981). People who are more knowledgeable about a topic are presumably better equipped to form relevant opinions and are less likely to offer NO responses (Faulkenberry & Mason 1978; Krosnick & Milburn 1990; Leigh & Martin 1987; Rapoport 1981, 1982). The more interested a person is in a topic, the more likely he or she is to form opinions on it, and the less likely he or she is to offer NO responses (Francis & Busch 1975; Krosnick & Milburn 1990; Norpoth & Buchanan 1992; Rapoport 1979, 1982; Wright & Niemi 1983). Opinion formation is presumably facilitated by exposure to information about a topic, and, in fact, greater exposure to the news media is associated with decreased NO answers to political opinion questions (Faulkenberry & Mason 1978, Krosnick & Milburn 1990, Wright & Niemi 1983). The more often a person performs behaviors that can be informed or shaped by an attitude, the more motivated that person is to form such an attitude, and the less likely that person is to say he or she has no opinion on an issue (Durand & Lambert 1988, Krosnick & Milburn 1990). The stronger a person's attitudes are, the less

likely he or she is to say “don’t know” when asked about their attitudes toward other objects in the domain (Wright & Niemi 1983). The greater an individual’s perception of his or her ability to process and understand information relevant to an attitude object, the less likely he or she is to say “don’t know” when asked about it (Krosnick & Milburn 1990). The more practical use a person believes there is in possessing attitudes toward an object, the less likely he or she is to say “don’t know” when asked to report such attitudes (Francis & Busch 1975; Krosnick & Milburn 1990). And people who consider a particular issue to be of less personal importance are more attracted to NO filters (Bishop et al 1980, Schuman & Presser 1981).

This suggests that NO options should increase the quality of data obtained by a questionnaire. By offering a NO option, respondents would be discouraged from offering meaningless opinions. Remarkably, this is not the case: offering a NO option does not increase the reliability of data obtained (Krosnick & Berent 1990, McClendon & Alwin 1993, Poe et al 1988). Associations between variables generally do not increase in strength when NO options are offered (Presser 1977, Sanchez & Morchio 1992, Schuman & Presser 1981), nor do answers become less susceptible to systematic measurement error caused by nonsubstantive aspects of question design (McClendon 1991). Asking people who offer NO responses to express an opinion anyhow leads to the expression of valid and predictive views (Gilljam & Granberg 1993, Visser et al 1999).

More evidence raises questions about the reliability of NO responses. The frequency of NO responses to a set of items is fairly consistent across different question sets in the same questionnaire (e.g. Cronbach 1950, Durand et al 1983, Durand & Lambert 1988, Fonda 1951, Leigh & Martin 1987, Lorge 1937) and over time (Krosnick & Milburn 1990, Rapoport 1982, Rosenberg et al 1955, Sigelman et al 1982). But there is a fair amount of random variation in whether a person expresses no opinion when answering any particular item (Butler & Stokes 1969, DuBois & Burns 1975, Durand et al 1983, Eisenberg & Wesman 1941, Lentz 1934). This random variation casts further doubt on the notion that NO responses genuinely, precisely, and comprehensively reflect lack of opinions.

Although NO responses sometimes occur because people have no information about an object, they occur more often for a variety of other reasons. People sometimes offer such responses because they feel ambivalent about the issue (e.g. Coombs & Coombs 1976, Klopfer & Madden 1980) or because they do not understand the meaning of a question or the answer choices (e.g. Converse 1976, Faulkenberry & Mason 1978, Fonda 1951, Klare 1950). Some NO responses occur because respondents think that they must know a lot about a topic to legitimately express an opinion (Berger & Sullivan 1970, Hippler & Schwarz 1989, McClendon 1986b), and some occur because people are avoid-

ing honestly answering a question in a way that would be unflattering (Cronbach 1950, Fonda 1951, Johanson et al 1993, Kahn & Hadley 1949, Rosenberg et al 1955). Some NO responses occur because interviewers expect that it will be difficult to administer items, and this expectation becomes a self-fulfilling prophecy (Singer et al 1983).

NO responses appear to result from satisficing as well (Krosnick 1991). According to this perspective, offering a NO option may discourage respondents from providing thoughtful answers. That is, respondents who are disposed to satisfice because of low ability to optimize, low motivation, or high task difficulty may be likely to select NO options as a way of avoiding the cognitive work necessary to generate an optimal answer. If a NO option is not offered, these respondents would be less likely to satisfice and might optimize instead.

Some of the evidence reviewed earlier is consistent with this reasoning. For example, NO filters attract respondents with limited cognitive skills. This is consistent with the notion that NO responses reflect satisficing caused by low cognitive skills. Also, NO responses are common among people for whom an issue is low in personal importance, of little interest, and arouses little affective involvement, and this may be because of lowered motivation to optimize under these conditions. Furthermore, people are likely to say they have no opinion when they feel they lack the ability to formulate informed opinions and when they feel there is little value in formulating such opinions. These associations may arise at the time of attitude measurement: Low motivation may inhibit a person from drawing on available knowledge to formulate and carefully report a substantive opinion on an issue. Also consistent with this perspective are demonstrations that NO responses become more common as questions become more difficult. Although all of this evidence is consistent with the notion that these responses reflect optimizing, it is also consistent with the satisficing view of NO responses.

Stronger support for the satisficing perspective comes from evidence that NO responses are more likely when questions appear later in a questionnaire, at which point motivation is waning (Culpepper et al 1992, Dickinson & Kirzner 1985, Ferber 1966, Ying 1989) and when respondents' intrinsic motivation to optimize has been undermined (Hansen 1980). NO responses are less common when the sponsor of a study is described as prestigious (Houston & Nevin 1977). Furthermore, inducements to optimize decrease NO responses (McDaniel & Rao 1980, Wotruba 1966).

SUMMARY The satisficing perspective offers new explanations for longstanding response patterns in questionnaire responses. The development of basic psychological theory in this fashion is a hallmark of the blossoming contemporary literature on survey methods.

CONCLUSION

The turn of the century provides an opportunity to reflect on the last 100 years and plot future courses of action in an informed way. Survey researchers are plotting their future with new visions of possibilities, because research is leading them to question old assumptions and to contemplate ways to improve their craft. The benefits of such efforts will be substantial both for psychologists who use survey methods as tools and for psychologists interested in understanding the workings of the human mind and the dynamics of social interaction.

ACKNOWLEDGMENT

The author thanks Catherine Heaney and Allyson Holbrook for helpful comments and Michael Tichy for heroic assistance in the manuscript preparation.

Visit the *Annual Reviews* home page at
<http://www.AnnualReviews.org>.

Literature Cited

- Abelson RP, Loftus EF, Greenwald AG. 1992. Attempts to improve the accuracy of self-reports of voting. In *Questions About Questions*, ed. JM Tanur, pp. 138–53. New York: Russell Sage
- Andrews FM. 1984. Construct validity and error components of survey measures: A structural modeling approach. *Public Opin. Q.* 48:409–42
- Atkinson RC, Shiffrin RM. 1968. Human memory: a proposed system and its control processes. In *The Psychology of Learning and Motivation: Advances in Research and Theory*, ed. KW Spence, JT Spence, 2:89–195. New York: Academic
- Ayidiya SA, McClendon M J. 1990. Response effects in mail surveys. *Public Opin. Q.* 54:229–47
- Babbie ER. 1990. *Survey Research Methods*. Belmont, CA: Wadsworth. 395 pp.
- Bachman JG, O'Malley PM. 1984. Yea-saying, nay-saying, and going to extremes: black-white differences in response styles. *Public Opin. Q.* 48:491–509
- Baddeley AD, Hitch GJ. 1977. Recency re-examined. In *Attention and Performance*, ed. S Dornic. Hillsdale, NJ: Erlbaum. Vol. 6.
- Bass BM. 1956. Development and evaluation of a scale for measuring social acquiescence. *J. Abnorm. Soc. Psychol.* 52: 296–99
- Bass BM. 1958. Famous sayings test: general manual. *Psychol. Rep.* 4:479–97
- Bass BM, Avolio BJ. 1989. Potential biases in leadership measures: How prototypes, leniency, and general satisfaction relate to ratings and rankings of transformational and transactional leadership constructs. *Educ. Psychol. Meas.* 49:509–27
- Becker SL. 1954. Why an order effect. *Public Opin. Q.* 18:271–78
- Berg IA, Rapaport GM. 1954. Response bias in an unstructured questionnaire. *J. Psychol.* 38:475–81
- Berger PK, Sullivan JE. 1970. Instructional set, interview context, and the incidence of “don’t know” responses. *J. Appl. Psychol.* 54:414–16
- Bickart B, Felcher EM. 1996. Expanding and enhancing the use of verbal protocols in survey research. In *Answering Questions*, ed. N Schwarz, S Sudman. San Francisco, CA: Jossey-Bass
- Bischooping K. 1989. An evaluation of interviewer debriefing in survey pretests. In *New Techniques for Pretesting Survey Questions*, ed. CF Cannell, L Oskenberg, FJ Fowler, G Kalton, K Bischooping. Ann Arbor, MI: Survey Res. Cent.

- Bishop GF. 1987. Experiments with the middle response alternative in survey questions. *Public Opin. Q.* 51:220-32
- Bishop GF, Hippler HJ, Schwarz N, Strack F. 1988. A comparison of response effects in self-administered and telephone surveys. In *Telephone Survey Methodology*, ed. RM Groves, PP Biemer, LE Lyberg, JT Massey, WL Nicholls II, J Waksberg, pp. 321-34. New York: Wiley
- Bishop GF, Oldendick RW, Tuchfarber AJ. 1980. Experiments in filtering political opinions. *Polit. Behav.* 2:339-69
- Bishop GF, Oldendick RW, Tuchfarber AJ. 1986. Opinions on fictitious issues: the pressure to answer survey questions. *Public Opin. Q.* 50:240-50
- Bogart L. 1972. *Silent Politics: Polls and the Awareness of Public Opinion*. New York: Wiley-Interscience
- Brady HE. 1990. Dimension analysis of ranking data. *Am. J. Polit. Sci.* 34:1017-48
- Brehm J. 1993. *The Phantom Respondents*. Ann Arbor: Univ. Mich. Press
- Briggs CL. 1986. *Learning How To Ask: A Sociolinguistic Appraisal of the Role of the Interview in Social Science Research*. Cambridge: Cambridge Univ. Press. 155 pp.
- Brown P, Levinson SC. 1987. *Politeness: Some Universals in Language Use*. New York: Cambridge Univ. Press. 345 pp.
- Butler D, Stokes D. 1969. *Political Change in Britain: Forces Shaping Electoral Choice*. New York: St. Martin's. 516 pp.
- Cacioppo JT, Petty RE, Feinstein JA, Jarvis WBG. 1996. Dispositional differences in cognitive motivation: the life and times of individuals varying in need for cognition. *Psychol. Bull.* 119:197-253
- Calsyn RJ, Roades LA, Calsyn DS. 1992. Acquiescence in needs assessment studies of the elderly. *The Gerontol.* 32: 246-52
- Campbell DT, Mohr PJ. 1950. The effect of ordinal position upon responses to items in a checklist. *J. Appl. Psychol.* 34:62-67
- Cannell CF, Miller PV, Oksenberg L. 1981. Research on interviewing techniques. In *Sociological Methodology*, ed. S Leinhardt, pp. 389-437. San Francisco, CA: Jossey-Bass
- Carp FM. 1974. Position effects on interview responses. *J. Gerontol.* 29:581-87
- Carpenter PA, Just MA. 1975. Sentence comprehension: a psycholinguistic processing model of verification. *Psychol. Rev.* 82: 45-73
- Carr LG. 1971. The role items and acquiescence. *Am. Sociol. Rev.* 36:287-93
- Chan JC. 1991. Response-order effects in Likert-type scales. *Educ. Psychol. Meas.* 51:531-40
- Cialdini RB. 1993. *Influence: Science and Practice*. New York: Harper Collins. 253 pp. 3rd ed.
- Clancy KJ, Wachslers RA. 1971. Positional effects in shared-cost surveys. *Public Opin. Q.* 35:258-65
- Clare ICH, Gudjonsson GH. 1993. Interrogative suggestibility, confabulation, and acquiescence in people with mild learning disabilities (mental handicap): implications for reliability during police interrogations. *Br. J. Clin. Psychol.* 32:295-301
- Clark HH. 1996. *Using Language*. New York: Cambridge Univ. Press. 432 pp.
- Clark HH, Chase WG. 1972. On the process of comparing sentences against pictures. *Cogn. Psychol.* 3:472-517
- Clark HH, Chase WG. 1974. Perceptual coding strategies in the formation and verification of descriptions. *Mem. Cogn.* 2: 101-11
- Clark HH, Clark EV. 1977. *Psychology and Language*. New York: Harcourt Brace Jovanovich. 608 pp.
- Clausen A. 1968. Response validity: vote report. *Public Opin. Q.* 32:588-606
- Coker MC, Knowles ES. 1987. *Testing alters the test scores: Test-retest improvements in anxiety also occur within a test*. Presented at the Midwest. Psychol. Assoc. Annu. Meet., Chicago
- Converse JM. 1976. Predicting no opinion in the polls. *Public Opin. Q.* 40:515-30
- Converse JM. 1987. *Survey Research in the United States: Roots and Emergence 1890-1960*. Berkeley, Los Angeles: Univ. Calif. Press
- Converse JM, Presser S. 1986. *Survey Questions: Handcrafting the Standardized Questionnaire*. Beverly Hills, CA: Sage. 80 pp.
- Converse PE. 1964. The nature of belief systems in the mass public. In *Ideology and Discontent*, ed. DE Apter, pp. 206-61. New York: Free Press
- Coombs CH, Coombs LC. 1976. "Don't know": item ambiguity or respondent uncertainty? *Public Opin. Q.* 40:497-514
- Costa PT, McCrae RR. 1988. From catalog to classification: Murray's needs and the five-factor model. *J. Pers. Soc. Psychol.* 55:258-65
- Costa PT, McCrae RR. 1995. Solid ground in the wetlands: a reply to Block. *J. Pers. Soc. Psychol.* 117:216-20
- Cronbach LJ. 1941. An experimental comparison of the multiple true-false and multiple-choice tests. *J. Educ. Psychol.* 32:533-43
- Cronbach LJ. 1950. Further evidence on re-

- sponse sets and test design. *Educ. Psychol. Meas.* 10:3-31
- Culpepper IJ, Smith WR, Krosnick JA. 1992. *The impact of question order on satisficing in surveys*. Presented at Midwest Psychol. Assoc. Annu. Meet., Chicago
- Davis JA, Smith TW. 1996. *General Social Surveys, 1972-1996: Cumulative Codebook*. Chicago: Natl. Opin. Res. Cent.
- DeMaio TJ, Rothgeb JM. 1996. Cognitive interviewing techniques: in the lab and in the field. In *Answering Questions*, ed. N Schwarz, S Sudman, pp. 177-96. San Francisco, CA: Jossey-Bass
- Dickinson JR, Kirzner E. 1985. Questionnaire item omission as a function of within-group question position. *J. Bus. Res.* 13: 71-75
- Dickinson TL, Zellinger PM. 1980. A comparison of the behaviorally anchored rating mixed standard scale formats. *J. Appl. Psychol.* 65:147-54
- DuBois B, Burns JA. 1975. An analysis of the meaning of the question mark response category in attitude scales. *Educ. Psychol. Meas.* 35:869-84
- Durand RM, Guffey HJ, Planchon JM. 1983. An examination of the random versus non-random nature of item omission. *J. Mark. Res.* 20:305-13
- Durand RM, Lambert ZV. 1988. Don't know responses in surveys: analyses and interpretational consequences. *J. Bus. Res.* 16: 169-88
- Ehrlich HL, Rinehart JW. 1965. A brief report on the methodology of stereotype research. *Soc. Forces* 43:564-75
- Eisenberg P, Wesman AG. 1941. Consistency in response and logical interpretation of psychoneurotic inventory items. *J. Educ. Psychol.* 32:321-38
- Elig TW, Frieze IH. 1979. Measuring causal attributions for success and failure. *J. Pers. Soc. Psychol.* 37:621-34
- Falzhzik AM, Jolson MA. 1974. Statement polarity in attitude studies. *J. Mark. Res.* 11:102-5
- Faulkenberry GD, Mason R. 1978. Characteristics of nonopinion and no opinion response groups. *Public Opin. Q.* 42:533-43
- Ferber R. 1966. Item nonresponse in a consumer survey. *Public Opin. Q.* 30:399-415
- Fonda CP. 1951. The nature and meaning of the Rorschach white space response. *J. Abnorm. Soc. Psychol.* 46:367-77
- Forehand GA. 1962. Relationships among response sets and cognitive behaviors. *Educ. Psychol. Meas.* 22:287-302
- Forsyth BH, Lessler JT. 1991. Cognitive laboratory methods: a taxonomy. In *Measurement Error in Surveys*, ed. P Biemer, R Groves, L Lyberg, N Mathiowetz, S Sudman, pp. 393-418. New York: Wiley
- Foster RJ. 1961. Acquiescent response set as a measure of acquiescence. *J. Abnorm. Soc. Psychol.* 63:155-60
- Foster RJ, Grigg AE. 1963. Acquiescent response set as a measure of acquiescence: further evidence. *J. Abnorm. Soc. Psychol.* 67:304-6
- Fowler FJ, Cannell CF. 1996. Using behavioral coding to identify cognitive problems with survey questions. In *Answering Questions*, ed. N Schwarz, S Sudman. San Francisco, CA: Jossey-Bass
- Fowler FJ Jr, Mangione TW. 1990. *Standardized Survey Interviewing*. Newbury Park, CA: Sage. 151 pp.
- Franck JD, Busch L. 1975. What we don't know about "I don't know's." *Public Opin. Q.* 34:207-18
- Gage NL, Leavitt GS, Stone GC. 1957. The psychological meaning of acquiescence set for authoritarianism. *J. Abnorm. Soc. Psychol.* 55:98-103
- Geer JG. 1988. What do open-ended questions measure? *Public Opin. Q.* 52:365-71
- Gergen KJ, Back KW. 1965. Communication in the interview and the disengaged respondent. *Public Opin. Q.* 30:385-98
- Gilbert DT, Krull DS, Malone PS. 1990. Unbelieving the unbelievable: some problems in the rejection of false information. *J. Pers. Soc. Psychol.* 59:601-13
- Gill SN. 1947. How do you stand on sin? *Tide* 14:72
- Gilljam M, Granberg D. 1993. Should we take don't know for an answer? *Public Opin. Q.* 57:348-57
- Goldberg LR. 1990. An alternative "description of personality": the big-five factor structure. *J. Pers. Soc. Psychol.* 59: 1216-29
- Goldstein KM, Blackman S. 1976. Cognitive complexity, maternal child rearing, and acquiescence. *Soc. Behav. Pers.* 4:97-103
- Gove WR, Geerken MR. 1977. Response bias in surveys of mental health: an empirical investigation. *Am. J. Sociol.* 82:1289-317
- Granberg G, Holmberg S. 1991. Self-reported turnout and voter validation. *Am. J. Polit. Sci.* 35:448-59
- Greenwald AG, Carnot CG, Beach R, Young B. 1987. Increasing voting behavior by asking people if they expect to vote. *J. Appl. Psychol.* 72:315-18
- Grice HP. 1975. Logic and conversation. In *Syntax and Semantics 3: Speech Acts*, ed. P Cole, JL Morgan, pp. 41-58. New York: Academic
- Gruber RE, Lehmann DR. 1983. The effect of omitting response tendency variables from

- regression models. In *1983 AMA Winter Educators Conference: Research Methods Causal Models in Marketing*, ed. WR Darden, KB Monroe, WR Dillon, pp. 131–36. Chicago: Am. Mark. Assoc.
- Gudjonsson GH. 1990. The relationship of intellectual skills to suggestibility, compliance and acquiescence. *Pers. Individ. Differ.* 11:227–31
- Hanley C. 1959. Responses to the wording of personality test items. *J. Consult. Psychol.* 23:261–65
- Hanley C. 1962. The “difficulty” of a personality inventory item. *Educ. Psychol. Meas.* 22:577–84
- Hansen RA. 1980. A self-perception interpretation of the effect of monetary and non-monetary incentives on mail survey respondent behavior. *J. Mark. Res.* 17:77–83
- Hartley EL. 1946. *Problems in Prejudice*. New York: Kings’ Crown. 124 pp.
- Hathaway SR, McKinley JC. 1940. A multi-phasic personality schedule (Minnesota): I. Construction of the schedule. *J. Psychol.* 10:249–54
- Hawkins DI, Coney KA. 1981. Uninformed response error in survey research. *J. Mark. Res.* 18:370–74
- Heaven PCL. 1983. Authoritarianism or acquiescence? South African findings. *J. Soc. Psychol.* 119:11–15
- Henry GT. 1990. *Practical Sampling*. Newbury Park, CA: Sage
- Herzog AR, Bachman JG. 1981. Effects of questionnaire length on response quality. *Public Opin. Q.* 45:549–59
- Hippler HJ, Schwarz N. 1989. “No-opinion” filters: a cognitive perspective. *Int. J. Public Opin. Res.* 1:77–87
- Holmes C. 1974. A statistical evaluation of rating scales. *J. Mark. Res. Soc.* 16:86–108
- Houston MJ, Nevin JR. 1977. The effects of source and appeal on mail survey response patterns. *J. Mark. Res.* 14:374–78
- Hui CH, Triandis HC. 1985. The instability of response sets. *Public Opin. Q.* 49:253–60
- Hurd AW. 1932. Comparisons of short answer and multiple choice tests covering identical subject content. *J. Educ. Psychol.* 26:28–30
- Husek TR. 1961. Acquiescence as a response set and as a personality characteristic. *Educ. Psychol. Meas.* 21:295–307
- Israel GD, Taylor CL. 1990. Can response order bias evaluations? *Eval. Program. Plan.* 13:365–71
- Jackson DN. 1959. Cognitive energy level, acquiescence, and authoritarianism. *J. Soc. Psychol.* 49:65–69
- Johanson GA, Gips CJ, Rich CE. 1993. If you can’t say something nice: a variation on the social desirability response set. *Eval. Rev.* 17:116–22
- Johnson JD. 1981. Effects of the order of presentation of evaluative dimensions for bipolar scales in four societies. *J. Soc. Psychol.* 113:21–27
- Jordan LA, Marcus AC, Reeder LG. 1980. Response styles in telephone and household interviewing: a field experiment. *Public Opin. Q.* 44:210–22
- Kahn DF, Hadley JM. 1949. Factors related to life insurance selling. *J. Appl. Psychol.* 33:132–40
- Kalton G, Collins M, Brook L. 1978. Experiments in wording opinion questions. *Appl. Stat.* 27:149–61
- Kish L. 1965. *Survey Sampling*. New York: Wiley. 634 pp.
- Klare GR. 1950. Understandability and indefinite answers to public opinion questions. *Int. J. Opin. Attitude Res.* 4:91–96
- Klayman J, Ha Y. 1987. Confirmation, disconfirmation, and information in hypothesis-testing. *Psychol. Rev.* 94:211–28
- Klockars AJ, Yamagishi M. 1988. The influence of labels and positions in rating scales. *J. Educ. Meas.* 25:85–96
- Klopfert FJ, Madden TM. 1980. The middlemost choice on attitude items: ambivalence, neutrality, or uncertainty. *Pers. Soc. Psychol. Bull.* 6:97–101
- Knowles ES. 1988. Item context effects on personality scales: measuring changes the measure. *J. Pers. Soc. Psychol.* 55:312–20
- Knowles ES, Cook DA, Neville JW. 1989a. *Assessing adjustment improves subsequent adjustment scores*. Presented at the Annu. Meet. Am. Psychol. Assoc., New Orleans, LA
- Knowles ES, Cook DA, Neville JW. 1989b. *Modifiers of context effects on personality tests: Verbal ability and need for cognition*. Presented at the Annu. Meet. Midwest. Psychol. Assoc., Chicago
- Koriat A, Lichtenstein S, Fischhoff B. 1980. Reasons for confidence. *J. Exp. Psychol.: Hum. Learn. Mem.* 6:107–18
- Kraut AI, Wolfson AD, Rothenberg A. 1975. Some effects of position on opinion survey items. *J. Appl. Psychol.* 60:774–76
- Krosnick JA. 1991. Response strategies for coping with the cognitive demands of attitude measures in surveys. *Appl. Cogn. Psychol.* 5:213–36
- Krosnick JA. 1992. The impact of cognitive sophistication and attitude importance on response order effects and question order effects. In *Order Effects in Social and Psy-*

- chological Research*, ed. N Schwarz, S Sudman, pp. 203–18. New York: Springer
- Krosnick JA, Alwin DF. 1987. An evaluation of a cognitive theory of response-order effects in survey measurement. *Public Opin. Q.* 51:201–19
- Krosnick JA, Alwin DF. 1988. A test of the form-resistant correlation hypothesis: ratings, rankings, and the measurement of values. *Public Opin. Q.* 52:526–38
- Krosnick JA, Berent MK. 1990. *The impact of verbal labeling of response alternatives and branching on attitude measurement reliability in surveys*. Presented at the Annu. Meet. Am. Assoc. Public Opin. Res., Lancaster, PA
- Krosnick JA, Berent MK. 1993. Comparisons of party identification and policy preferences: the impact of survey question format. *Am. J. Polit. Sci.* 37:941–64
- Krosnick JA, Fabrigar LR. 1998. *Designing Good Questionnaires: Insights from Psychology*. New York: Oxford Univ. Press. In press
- Krosnick JA, Li F, Lehman DR. 1990. Conversational conventions, order of information acquisition, and the effect of base rates and individuating information on social judgments. *J. Pers. Soc. Psychol.* 59: 1140–52
- Krosnick JA, Milburn MA. 1990. Psychological determinants of political opinionation. *Soc. Cogn.* 8:49–72
- Krosnick JA, Narayan S, Smith WR. 1996. Satisficing in surveys: initial evidence. *New Direct. Eval.* 70:29–44
- Krosnick JA, Schuman H. 1988. Attitude intensity, importance, and certainty and susceptibility to response effects. *J. Pers. Soc. Psychol.* 54:940–52
- Kuethle JL. 1959. The positive response set as related to task performance. *J. Pers.* 27: 87–95
- Kunda Z, Fong GT, Sanitioso R, Reber E. 1993. Directional questions direct self-conceptions. *J. Exp. Soc. Psychol.* 29: 63–86
- Laumann EO, Michael RT, Gagnon JH, Michaels S. 1994. *The Social Organization of Sexuality: Sexual Practices in the United States*. Chicago: Univ. Chicago Press. 718 pp.
- Lavrakas PJ. 1993. *Telephone Survey Methods: Sampling, Selection, and Supervision*. Newbury Park, CA: Sage, 157 pp. 2nd ed.
- Leech GN. 1983. *Principles of Pragmatics*. London/New York: Longman. 250 pp.
- Leigh JH, Martin CR Jr. 1987. “Don’t know” item nonresponse in a telephone survey: effects of question form and respondent characteristics. *J. Mark. Res.* 24:418–24
- Lenski GE, Leggett JC. 1960. Caste, class, and deference in the research interview. *Am. J. Sociol.* 65:463–67
- Lentz TF. 1934. Reliability of the opinioaire technique studies intensively by the retest method. *J. Soc. Psychol.* 5:338–64
- Lorge I. 1937. Gen-like: Halo or reality. *Psychol. Bull.* 34:545–46
- Mathews CO. 1927. The effect of position of printed response words upon children’s answers to questions in two-response types of tests. *J. Educ. Psychol.* 18:445–57
- Mazmanian D, Mendonca JD, Holden RR, Dufton B. 1987. Psychopathology and response styles in the SCL-90 responses of acutely distressed persons. *J. Psychopathol. Behav. Assess.* 9:135–48
- McClendon MJ. 1986a. Response-order effects for dichotomous questions. *Soc. Sci. Q.* 67:205–11
- McClendon MJ. 1986b. Unanticipated effects of no opinion filters on attitudes and attitude strength. *Soc. Perspect.* 29:379–95
- McClendon MJ. 1991. Acquiescence and recency response-order effects in interview surveys. *Soc. Methods Res.* 20:60–103
- McClendon MJ, Alwin DF. 1993. No-opinion filters and attitude measurement reliability. *Soc. Methods Res.* 21:438–64
- McDaniel SW, Rao CP. 1980. The effect of monetary inducement on mailed questionnaire response quality. *J. Mark. Res.* 17: 265–68
- McIntyre SH, Ryans AB. 1977. Time and accuracy measures for alternative multidimensional scaling data collection methods: some additional results. *J. Mark. Res.* 14:607–10
- Messick S, Frederiksen N. 1958. Ability, acquiescence, and “authoritarianism.” *Psychol. Rep.* 4:687–97
- Miethe TD. 1985. The validity and reliability of value measurements. *J. Pers.* 119: 441–53
- Miller N, Campbell DT. 1959. Recency and primacy in persuasion as a function of the timing of speeches and measurement. *J. Abnorm. Soc. Psychol.* 59:1–9
- Mingay DJ, Greenwell MT. 1989. Memory bias and response-order effects. *J. Off. Stat.* 5:253–63
- Mirowsky J, Ross CE. 1991. Eliminating defense and agreement bias from measures of the sense of control: a 2 × 2 index. *Soc. Psychol. Q.* 54:127–45
- Mishler EG. 1986. *Research Interviewing*. Cambridge, MA: Harvard Univ. Press. 189 pp.
- Mosteller F, Hyman H, McCarthy PJ, Marks ES, Truman DB. 1949. *The Pre-Election Polls of 1948: Report to the Committee on*

- Analysis of Pre-Election Polls and Forecasts*. New York: Soc. Sci. Res. Council.
- Munson JM, McIntyre SH. 1979. Developing practical procedures for the measurement of personal values in cross-cultural marketing. *J. Mark. Res.* 16:48–52
- Narayan S, Krosnick JA. 1996. Education moderates some response effects in attitude measurement. *Public Opin. Q.* 60: 58–88
- Nathan BR, Alexander RA. 1985. The role of inferential accuracy in performance rating. *Acad. Manage. Rev.* 10:109–15
- Neidell LA. 1972. Procedures for obtaining similarities data. *J. Mark. Res.* 9:335–37
- Nelson D. 1985. Informal testing as a means of questionnaire development. *J. Off. Stat.* 1: 79–88
- Norpoth H, Buchanan B. 1992. Wanted: the education president: issue trespassing by political candidates. *Public Opin. Q.* 56: 87–99
- Parten M. 1950. *Surveys, Polls, and Samples: Practical Procedures*. New York: Harper. 624 pp.
- Payne JD. 1971. The effects of reversing the order of verbal rating scales in a postal survey. *J. Mark. Res. Soc.* 14:30–44
- Payne SL. 1949/1950. Case study in question complexity. *Public Opin. Q.* 13:653–58
- Payne SL. 1950. Thoughts about meaningless questions. *Public Opin. Q.* 14:687–96
- Peters DL, McCormick EJ. 1966. Comparative reliability of numerically anchored versus job-task anchored rating scales. *J. Appl. Psychol.* 50:92–96
- Pew Research Center. 1998. *Opinion poll experiment reveals conservative opinions not underestimated, but racial hostility missed*. Internet posting, <http://www.people-press.org/resprpt.htm>, March 27
- Poe GS, Seeman I, McLaughlin J, Mehl E, Dietz M. 1988. Don't know boxes in factual questions in a mail questionnaire. *Public Opin. Q.* 52:212–22
- Presser S. 1977. *Survey question wording and attitudes in the general public*. PhD thesis. Univ. Mich, Ann Arbor. 370 pp.
- Presser S. 1990. Measurement issues in the study of social change. *Soc. Forces* 68: 856–68
- Presser S, Blair J. 1994. Do different methods produce different results? In *Sociological Methodology*, ed. PV Marsden, pp. 73–104. Cambridge, MA: Blackwell
- Presser S, Traugott MW, Traugott S. 1990. *Vote "over" reporting in surveys: the records or the respondents?* Presented at Int. Conf. Measure. Errors, Tucson, AZ
- Quinn SB, Belson WA. 1969. *The Effects of Reversing the Order of Presentation of Verbal Rating Scales in Survey Interviews*. London: Survey Res. Cent.
- Rankin WL, Grube JW. 1980. A comparison of ranking and rating procedures for value system measurement. *Eur. J. Soc. Psychol.* 10:233–46
- Rapoport RB. 1979. What they don't know can hurt you. *Am. J. Polit. Sci.* 23:805–15
- Rapoport RB. 1981. The sex gap in political persuading: Where the "structuring principle" works. *Am. J. Polit. Sci.* 25:32–48
- Rapoport RB. 1982. Sex differences in attitude expression: a generational explanation. *Public Opin. Q.* 46:86–96
- Remmers HH, Marschat LE, Brown A, Chapman I. 1923. An experimental study of the relative difficulty of true-false, multiple-choice, and incomplete-sentence types of examination questions. *J. Educ. Psychol.* 14:367–72
- Reynolds TJ, Jolly JP. 1980. Measuring personal values: an evaluation of alternative methods. *J. Mark. Res.* 17:531–36
- Roberts RT, Forthofer RN, Fabrega H. 1976. The Langer items and acquiescence. *Soc. Sci. Med.* 10:69–75
- Robinson CE. 1932. *Straw Votes*. New York: Columbia Univ. Press. 203 pp.
- Robinson JP, Shaver PR, Wrightsman LS. 1991. *Measures of Personality and Social Psychological Attitudes*. San Diego, CA: Academic. 735 pp.
- Rosenberg N, Izard CE, Hollander EP. 1955. Middle category response: reliability and relationship to personality and intelligence variables. *Educ. Psychol. Meas.* 15:281–90
- Ross CE, Mirowsky J. 1984. Socially-desirable response and acquiescence in a cross-cultural survey of mental health. *J. Health Soc. Behav.* 25:189–97
- Ross CK, Steward CA, Sinacore JM. 1995. A comparative study of seven measures of patient satisfaction. *Med. Care* 33: 392–406
- Sanchez ME, Morchio G. 1992. Probing "don't know" answers. *Public Opin. Q.* 56:454–74
- Schober MF, Conrad FG. 1997. Does conversational interviewing reduce survey measurement error? *Public Opin. Q.* 61: 576–602
- Schriesheim CA, Hinkin TR, Podsakoff PM. 1991. Can ipsative and single-item measures produce erroneous results in field studies of French and Raven's 1959 five bases of power? An empirical investigation. *J. Appl. Psychol.* 76:106–14
- Schuman H, Ludwig J, Krosnick JA. 1986. The perceived threat of nuclear war, salience, and open questions. *Public Opin. Q.* 50:519–36

- Schuman H, Presser S. 1981. *Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording, and Context*. New York: Academic. 370 pp.
- Schutz RE, Foster RJ. 1963. A factor analytic study of acquiescent and extreme response set. *Educ. Psychol. Meas.* 23: 435-47
- Schwarz N. 1996. *Cognition and Communication: Judgmental Biases, Research Methods, and the Logic of Conversation*. Mahwah, NJ: Erlbaum
- Schwarz N, Hippler HJ. 1991. Response alternatives: the impact of their choice and presentation order. In *Measurement Error in Surveys*, ed. P Biemer, RM Groves, LE Lyberg, NA Mathiowetz, S Sudman, pp. 41-56. New York: Wiley
- Schwarz N, Hippler HJ, Deutsch B, Strack F. 1985. Response scales: effects of category range on reported behavior and subsequent judgments. *Public Opin. Q.* 49: 388-95
- Schwarz N, Hippler HJ, Noelle-Neumann E. 1992. A cognitive model of response-order effects in survey measurement. In *Context Effects in Social and Psychological Research*, ed. N Schwarz, S Sudman, New York: Springer-Verlag
- Schwarz N, Knauper B, Hippler HJ, Noelle-Neumann E, Clark LF. 1991. Rating scales: Numeric values may change the meaning of scale labels. *Public Opin. Q.* 55:570-82
- Schwarz N, Strack F. 1985. Cognitive and affective processes in judgments of subjective well-being: a preliminary model. In *Economic Psychology*, ed. H Brandstatter, E Kirchler, pp. 439-47. Linz, Austria: R. Tauner
- Shaw ME, Wright JM. 1967. *Scales for the Measurement of Attitudes*. New York: McGraw-Hill. 604 pp.
- Sherif CW, Sherif M, Nebergall RE. 1965. *Attitude and Attitude Change*. Philadelphia: Saunders. 264 pp.
- Sherif M, Hovland CI. 1961. *Social Judgment: Assimilation and Contrast Effects in Communication and Attitude Change*. New Haven, CT: Yale Univ. Press
- Sigelman CK, Winer JL, Schoenrock CJ. 1982. The responsiveness of mentally retarded persons to questions. *Educ. Train. Mental. Retard.* 17:120-24
- Sigelman L. 1981. Question-order effects on presidential popularity. *Public Opin. Q.* 45:199-207
- Sigelman L. 1982. The nonvoting voter in voting research. *Am. J. Polit. Sci.* 26: 47-56
- Silk AJ. 1971. Response set and the measurement of self-designated opinion leadership. *Public Opin. Q.* 35:383-97
- Silver BD, Anderson BA, Abramson RP. 1986. Who overreports voting? *Am. Polit. Sci. Rev.* 80:613-24
- Simon HA. 1957. *Models of Man*. New York: Wiley. 287 pp.
- Singer E, Frankel MR, Glassman MB. 1983. The effect of interviewer characteristics and expectations on response. *Public Opin. Q.* 47:68-83
- Small DO, Campbell DT. 1960. The effect of acquiescence response-set upon the relationship of the F scale and conformity. *Sociometry* 23:69-71
- Smith TW. 1983. The hidden 25 percent: an analysis of nonresponse in the 1980 General Social Survey. *Public Opin. Q.* 47: 386-404
- Smith TW. 1987. That which we call welfare by any other name would smell sweeter: an analysis of the impact of question wording on response patterns. *Public Opin. Q.* 51:75-83
- Sniderman P, Grob DB. 1996. Innovations in experimental design in attitude surveys. *Annu. Rev. Sociol.* 22:377-400
- Steeh C. 1981. Trends in nonresponse rates. *Public Opin. Q.* 45:40-57
- Suchman L, Jordan B. 1990. Interactional troubles in face-to-face survey interviews. *J. Am. Stat. Assoc.* 85:232-53
- Suchman L, Jordan B. 1992. Validity and the collaborative construction of meaning in face-to-face surveys. In *Questions About Questions*, ed. J Tanur, pp. 241-67. New York: Russell Sage Found.
- Sudman S, Bradburn NM, Schwarz N. 1996. *Thinking about Answers: The Application of Cognitive Processes to Survey Methodology*. San Francisco, CA: Jossey-Bass. 304 pp.
- Taylor JR, Kinnear TC. 1971. Empirical comparison of alternative methods for collecting proximity judgments. *Am. Market. Assoc. Proc. Fall Conf.*, pp. 547-50
- Tourangeau R, Rasinski KA. 1988. Cognitive processes underlying context effects in attitude measurement. *Psychol. Bull.* 103: 299-314
- Tourangeau R, Rips L, Rasinski K. 1998. *The Psychology of Survey Response*. New York: Cambridge Univ. Press. In press
- Traugott MW, Groves RM, Lepkowski JM. 1987. Using dual frame designs to reduce nonresponse in telephone surveys. *Public Opin. Q.* 51:522-39
- Traugott MW, Katosh JP. 1979. Response validity in surveys of voting behavior. *Public Opin. Q.* 43:359-77
- Trott DM, Jackson DN. 1967. An experimen-

- tal analysis of acquiescence. *J. Exp. Res. Pers.* 2:278-88
- Vaillancourt PM. 1973. Stability of children's survey responses. *Public Opin. Q.* 37: 373-87
- Visser PS, Krosnick JA, Marquette J, Curtin M. 1996. Mail surveys for election forecasting? An evaluation of the Columbus Dispatch poll. *Public Opin. Q.* 60:181-227
- Visser PS, Krosnick JA, Marquette J, Curtin M. 1999. Improving election forecasting: allocation of undecided respondents, identification of likely voters, and response order effects. In *Election Polls, the News Media, and Democracy*, ed. P Lavrakas, M Traugott. In press
- Warwick DP, Lininger CA. 1975. *The Sample Survey: Theory and Practice*. New York: McGraw-Hill. 344 pp.
- Webster H. 1958. Correcting personality scales for response sets or suppression effects. *Psychol. Bull.* 55:62-64
- Weisberg HF, Krosnick JA, Bowen BD. 1996. *An Introduction to Survey Research, Polling, and Data Analysis*. Newbury Park, CA: Sage. 394 pp. 3rd ed.
- Wotruba TR. 1966. Monetary inducements and mail questionnaire response. *J. Mark. Res.* 3:398-400
- Wright JR, Niemi RG. 1983. Perceptions of issue positions. *Polit. Behav.* 5:209-23
- Yalch RF. 1976. Pre-election interview effects on voter turnout. *Public Opin. Q.* 40: 331-36
- Ying Y. 1989. Nonresponse on the center for epidemiological studies-depression scale in Chinese Americans. *Int. J. Soc. Psychol.* 35:156-63
- Yzerbyt VY, Leyens J. 1991. Requesting information to form an impression: the influence of valence and confirmatory status. *J. Exp. Soc. Psychol.* 27:337-56
- Zaller J, Feldman S. 1992. A simple theory of the survey response: answering questions versus revealing preferences. *Am. J. Polit. Sci.* 36:579-616
- Zuckerman M, Bernieri F, Koestner R, Rosenthal R. 1989. To predict some of the people some of the time: in search of moderators. *J. Pers. Soc. Psychol.* 57:279-93