

# 3

## Attitude Measurement

NORBERT SCHWARZ

*Institute for Social Research, University of Michigan*

As currently used in psychology, the term *attitude* refers to a hypothetical construct, namely a predisposition to evaluate some object in a favorable or unfavorable manner (Eagly & Chaiken, 1993; Prislin & Crano, this volume). This predisposition cannot be directly observed and needs to be inferred from individuals' responses to the attitude object. These responses can run from overt behavior (such as approaching or avoiding the object) and explicit verbal statements (e.g., answers to an attitude question) to covert responses, which may be outside of the person's awareness (such as minute facial expressions or the speed with which a letter string can be recognized as a meaningful word). In principle, any one of these responses can be used to infer a person's attitude; however, each response may be influenced by variables other than the person's evaluative predisposition toward the attitude object, raising complex theoretical issues. Moreover, the same person's responses to different attitude measures may suggest different underlying attitudes; for example, a person's verbal statements may not converge with the person's overt behavior or spontaneous facial expressions.

This chapter provides an introduction to the most commonly used measurement procedures; it is organized as follows. The first section addresses direct attitude questions. It reviews the cognitive and communicative processes involved in answering attitude questions and discusses common question and response scale formats. As will become apparent, respondents' answers to attitude questions are highly context dependent and researchers have developed a number of alternative procedures, often in the hope that they would provide a less context dependent assessment of attitudes. The second section reviews some of these recent "implicit" attitude measures, most of which are based on variants of response time measurement (see also Devos, this volume).

A growing body of research indicates that these implicit measures are just as context dependent as explicit attitude reports. Alternatively, researchers can rely on psychophysiological measures or on behavioral observation, reviewed in the

third section. The chapter concludes with a discussion of the theoretical implications of the observed context dependency of attitude measurement by juxtaposing the traditional emphasis on evaluative predispositions with the alternative view that attitudes are context sensitive evaluations, constructed on the spot.

## DIRECT QUESTIONS: EXPLICIT SELF-REPORTS OF ATTITUDES

Most researchers rely on respondents' answers to direct attitude questions, like, "Do you approve or disapprove of how President Bush is handling his job?" Direct questions are the most feasible procedure for assessing the attitudes of the population at large, as is done in representative sample surveys. In laboratory research, direct questions can be supplemented with more indirect procedures, like psychophysiological or response time measurement. The use of direct questions is based on the premise that people have introspective access to their attitudes and are aware of what they like and dislike (for a discussion see Strack & Schwarz, 2007), whereas most other attitude measures do not require this assumption.

As attitude researchers have known for many decades, self-reports of attitudes are highly context dependent and minor changes in question wording, question format, and question order can profoundly affect the obtained results (for early reviews see Cantril, 1944; Payne, 1951). Since the early 1980s, psychologists and survey methodologists investigated the underlying cognitive and communicative processes and this section summarizes what has been learned. Sudman, Bradburn, and Schwarz (1996) and Tourangeau, Rips, and Rasinski (2000) provide more comprehensive reviews of this work, including self-reports of attitudes as well as behaviors (see also Schwarz & Oyserman, 2001).

### *Respondents' Tasks*

Answering an attitude question involves several tasks. As a first step, respondents need to understand the question to determine which information they are to provide. Next, they need to retrieve relevant information from memory to form an attitude judgment. In most cases, they cannot report this judgment in their own words, but need to format their answer to fit the response alternatives provided by the researcher. Moreover, they may want to edit their judgment before they report it, due to reasons of social desirability and self-presentation. Accordingly, comprehension, retrieval, judgment, formatting, and editing are the key components of the response process (see Strack & Martin, 1987; Sudman et al., 1996, chapter 3, for more detailed discussions). Performance at each of these steps is strongly influenced by contextual features.

### *Question Comprehension*

The key issue at the comprehension stage is whether the respondent's understanding of the question matches what the researcher had in mind: Is the attitude object

that the respondent reports on the one that the researcher intended? Does the respondent's understanding tap the intended facet of the issue and the intended evaluative dimension? Not surprisingly, researchers are urged to write clear and simple questions and to avoid unfamiliar or ambiguous terms. Bradburn, Sudman, and Wansink (2004) provide excellent advice in this regard. Even familiar terms, however, are open to interpretation and respondents draw on the context of the question to infer which meaning the researcher has in mind. Hence, the term *drugs* acquires a different meaning in the context of a health insurance survey than in the context of a crime survey. This use of contextual information is licensed by the tacit norms that underlie the conduct of conversation in daily life (Grice, 1975), where listeners are expected to take the content of preceding utterances into account when they interpret the next one. Research participants bring these conversational norms to the research situation and assume that all contributions of the researcher are relevant to the ongoing "conversation" (for reviews see Clark & Schober, 1992; Schwarz, 1994, 1996). These contributions include the study introduction and the content of preceding questions, as well as the specific wording of the question and many apparently "formal" features of the questionnaire.

Suppose, for example, that respondents are asked to report how successful they have been in life, along an 11-point rating scale, ranging from "not so successful" to "extremely successful." To provide a rating, they need to determine what "not so successful" means: Does it refer to the absence of outstanding achievements or to the presence of failure? To infer the intended meaning they may draw on formal characteristics of the rating scale. When the numeric values of the scale run from 0 = "not so successful" to 10 = "extremely successful," respondents interpret "not so successful" as pertaining to the absence of noteworthy achievements; but when the scale runs from -5 = "not so successful" to +5 = "extremely successful," they interpret "not so successful" as pertaining to the presence of failure. Because people are more likely to lack great achievements than to experience great failures, these differences in interpretation result in dramatic shifts in the obtained ratings. Specifically, 34% of a German sample endorsed a value between 0 and 5 on the 0 to 10 scale, whereas only 13% endorsed a formally equivalent value between -5 and 0 on the -5 to +5 scale (Schwarz, Knäuper, Hippler, Noelle-Neumann, & Clark, 1991). In general, a minus-to-plus rating scale format conveys that the researcher has a bipolar dimension in mind, where one endpoint refers to the opposite of the other. In contrast, a format that presents only positive numbers conveys that the researcher has a unipolar dimension in mind, where the numbers pertain to different degrees of the presence of the same attribute. Schwarz (1994, 1996) has reviewed what respondents infer from different elements of a questionnaire and the procedures used in laboratory experiments, and highlights how their inferences are consistent with normal conversational conduct in everyday life.

To safeguard against unintended question interpretations, researchers developed cognitive interviewing techniques to assess respondents' interpretation of questions at the pretest stage. Willis (2004) and the contributions in Schwarz and Sudman (1996) review these methods and provide advice on their use. Given the context dependency of question interpretations, these techniques should not be applied to isolated questions; instead, the question needs to be presented in the

context in which it will be used in the actual study. When properly employed, cognitive interviewing at the questionnaire development stage, and sensitive revisions, can ensure that respondents understand the final question as intended.

### *Information Retrieval and Judgment*

Once respondents determine what the question refers to, they need to recall relevant information from memory. In some cases, they may have direct access to a previously formed judgment that they can offer as an answer. In most cases, however, they will not find an appropriate answer readily stored in memory and will need to develop a judgment on the spot. To do so, they need to form a mental representation of the attitude object, and of a standard against which the object is evaluated. The resulting judgment depends on which information happens to come to mind at that point in time and on how this information is used.

As a large body of social cognition research demonstrates, people rarely retrieve all information that may bear on an attitude object; rather, they truncate the search process as soon as enough information has come to mind to form a judgment with sufficient subjective certainty (for reviews see Bodenhausen & Wyer, 1987; Wyer & Srull, 1989). Hence, the judgment is disproportionately influenced by the first few pieces of information that come to mind. Whereas some information may always come to mind when the person thinks of a particular object (and is therefore called chronically accessible), other information may be only temporarily accessible, for example, because it has been brought to mind by preceding questions. Changes in what is temporarily accessible are at the heart of many context effects in attitude measurement, including question order and response order effects, as discussed below. In contrast, chronically accessible information contributes some stability to attitude judgments.

**Question Order Effects** How accessible information influences the judgment depends on how it is used (Schwarz & Bless, 1992, 2007). Information that is included in the temporary representation formed of the target results in assimilation effects; that is, the inclusion of positive (negative) information results in a more positive (negative) judgment. The size of assimilation effects increases with the amount and extremity of temporarily accessible information and decreases with the amount and extremity of chronically accessible information included in the representation of the target (Bless, Schwarz, & Wänke, 2003). For example, Schwarz, Strack, and Mai (1991) asked respondents to report their marital satisfaction and their general life-satisfaction in different question orders. When the general life-satisfaction question was asked first, it correlated with marital satisfaction  $r = .32$ . Reversing the question order, however, increased this correlation to  $r = .67$ . This reflects that the marital satisfaction question brought marriage related information to mind, and respondents included the information in the representation formed of their lives in general.

This increase in correlation was attenuated ( $r = .43$ ) when questions about three different life-domains (job, leisure time, and marriage) preceded the general question, thus bringing a more diverse range of information to mind. Parallel influences

were observed in the mean reports. Happily married respondents reported higher, and unhappily married respondents reported lower, general life-satisfaction when their attention was drawn to their marriage by the preceding question. However, the same piece of accessible information that may elicit an assimilation effect may also result in a contrast effect; that is, in a more negative (positive) judgment, the more positive (negative) information is brought to mind. This is the case when the information is excluded from, rather than included in, the cognitive representation formed of the target (Schwarz & Bless, 1992, 2007). As a first possibility, suppose that a given piece of information with positive (negative) implications is excluded from the representation of the target category. If so, the representation will contain less positive (negative) information, resulting in less positive (negative) judgments. This possibility is referred to as a subtraction based contrast effect. The size of subtraction based contrast effects increases with the amount and extremity of the temporarily accessible information that is excluded from the representation of the target, and decreases with the amount and extremity of the information that remains in the representation of the target.

For example, the above study (Schwarz et al., 1991) included a condition in which the marital satisfaction and life-satisfaction questions were introduced with a joint lead-in that read, "We now have two questions about your life. The first pertains to your marriage and the second to your life in general." This lead-in was designed to evoke the conversational maxim of quantity (Grice, 1975), which enjoins speakers to avoid redundancy when answering related questions. Accordingly, respondents who had just reported on their marriage should now disregard this aspect of their lives when answering the general life-satisfaction question. Confirming this prediction, happily married respondents now reported lower general life-satisfaction, whereas unhappily married respondents reported higher life-satisfaction, indicating that they excluded the positive (negative) marital information from the representation formed of their lives in general. These diverging effects reduced the correlation to  $r = .18$ , from  $r = .67$  when the same questions were asked in the same order without a joint lead-in. Finally, a control condition in which the general life-satisfaction question was reworded to, "Aside from your marriage, which you already told us about, how satisfied are you with your life in general?" resulted in a similarly low correlation of  $r = .20$ . Such subtraction based contrast effects are limited to the specific target (here, one's life in general), reflecting that merely "subtracting" a piece of information (here, one's marriage) does only affect this specific representation.

As a second possibility, respondents may not only exclude accessible information from the representation formed of the target, but may also use this information in constructing a standard of comparison or scale anchor. If the implications of the temporarily accessible information are more extreme than the implications of the chronically accessible information used in constructing a standard, they result in a more extreme standard, eliciting contrast effects for that reason. The size of comparison based contrast effects increases with the extremity and amount of temporarily accessible information used in constructing the standard, and decreases with the amount and extremity of chronically accessible information used in making this construction. In contrast to subtraction based comparison effects, which

are limited to a specific target, comparison based contrast effects generalize to all targets to which the standard is applicable.

As an example, consider the impact of political scandals on assessments of the trustworthiness of politicians. Not surprisingly, thinking about a politician who was involved in a scandal, say Richard Nixon, decreases trust in politicians in general. This reflects that the exemplar is included in the representation formed of the target “politicians in general.” If the trustworthiness question pertains to a specific politician, however, say Bill Clinton, the primed exemplar cannot be included in the representation formed of the target—after all, Bill Clinton is not Richard Nixon. In this case, Richard Nixon may serve as a standard of comparison, relative to which Bill Clinton seems more trustworthy than would otherwise be the case. An experiment with German exemplars confirmed these predictions (Schwarz & Bless, 1992b): Thinking about a politician who was involved in a scandal decreased the trustworthiness of politicians in general, but increased the trustworthiness of all specific exemplars assessed.

If a given piece of information is used in constructing a representation of the attitude object (resulting in assimilation effects), or of a standard of comparison (resulting in contrast effects), depends on a host of different variables, which are beyond the scope of this chapter. Schwarz and Bless (2007) and Sudman et al. (1996) review these variables and present a theoretical model that predicts the direction, size, and generalization of question order effects in attitude measurement.

**Response Order Effects** Respondents’ judgments are also influenced by the order in which response alternatives are presented within a question. To understand the underlying processes, suppose you are asked to provide a few good reasons why “divorce should be easier to obtain.” You can easily do so, but you could just as easily provide some reasons why “divorce should be more difficult to obtain.” When such alternatives are juxtaposed within a question (as in “Should divorce be easier to obtain or more difficult to obtain?”), the outcome depends on which alternative is considered first. When respondents first consider “easier” and generate some supportive thoughts, they are likely to truncate the search process and endorse this response option; but had they considered “more difficult,” the same process would have resulted in an endorsement of that option. Again, respondents’ judgments are based on the temporary representation formed of the attitude object (“divorce”), which is a function of the thoughts brought to mind by the response option considered first.

Which option respondents consider first depends on the order and mode in which the response alternatives are presented (Krosnick & Alwin, 1987). When presented in writing, respondents read down the list of response alternatives and elaborate on their implications in the order in which they are presented. In this mode, an alternative that elicits supporting thoughts is more likely to be endorsed when presented early rather than late on the list, giving rise to primacy effects. In contrast, when the alternatives are read to respondents, their opportunity to think about the early ones is limited by the need to listen to the later ones. In this case, they are more likely to work backwards, thinking first about the last alternative read to them, which is still “in their ears.” When the last alternative heard elicits

supporting thoughts, it is likely to be endorsed, giving rise to recency effects. As a result, a given alternative is more likely to be endorsed when presented early rather than late in a visual format (primacy effect), but when presented late rather than early in an auditory format (recency effect). Sudman et al. (1996) review these processes in more detail.

Response order effects are most likely to be obtained when respondents can generate supporting thoughts for several of the response alternatives presented to them, as in the above divorce example. When one alternative is attractive and the other unattractive, the order in which they are presented is unlikely to make a difference (Sudman et al., 1996). Finally, response order effects are more pronounced for older and less educated respondents (see Knäuper, 1999, for a meta-analysis), whose limited cognitive resources further enhance the focus on a single response alternative. This age-sensitivity of response order effects can invite misleading conclusions about cohort differences in the reported attitude, suggesting, for example, that older respondents are more liberal than younger respondents under one order condition, but more conservative under the other (see Schwarz & Knäuper, 2000).

### *Response Formatting*

Once respondents have formed a judgment, they can only report it in their own words when an open response format is used. Because open answers require cumbersome and expensive coding prior to statistical analysis, open response formats are rarely used in practice. Instead, respondents are usually asked to provide an answer in a closed response format, either by rating the attitude object along a scale or by selecting one of several substantive response alternatives presented to them.

**Categorical Response Alternatives** When the question offers several distinct opinions and asks the respondent to select the one that is closest to his or her own position, it is important to ensure that the set of response alternatives offered covers the whole range of plausible positions. Any opinion omitted from the list is unlikely to be reported, even when respondents are offered a general “other” response option, which they rarely use. For example, Schuman and Presser (1981) asked respondents what they consider “the most important thing for children to prepare them for life.” When the answer, “To think for themselves” was offered as part of a list, 61.5% of a representative sample endorsed it—yet only 4.6% volunteered an answer that could be assigned to this category when an open response format was used (Schuman & Presser, 1981). Such discrepancies reflect that the response alternatives clarify what the researcher is interested in and remind respondents of aspects they might otherwise not consider. Similarly, few respondents report not having an opinion on an issue when this option is not explicitly provided—yet, they may be happy to report so when “Don’t know” is offered as an alternative. Throughout, respondents work within the constraints imposed by the question (see Krosnick & Fabrigar, in press, for a review). In addition, respondents’ judgments are influenced by the order in which response alternatives are presented, as already seen.

**Rating Scales** Rating scales are the most commonly used response format in attitude measurement. Typically, a numerical scale with verbally labeled endpoints (e.g., -3 = strongly disagree; +3 = strongly agree) is presented and respondents are asked to check the number that best represents their opinion. As noted, the numeric values may themselves influence the interpretation of the verbal endpoints. Alternatively, each point of the rating scale may be labeled, a format that is more commonly used in telephone interviews than in self-administered questionnaires or laboratory experiments. In general, the retest reliability of fully labeled scales is somewhat higher than that of partially labeled scales. Moreover, retest reliability decreases as the number of scale points increases beyond seven, reflecting the difficulty of making many fine-grained distinctions. Krosnick and Fabrigar (in press) provide a comprehensive review of the relevant literature.

Respondents' use of rating scales is highly context dependent. As numerous studies demonstrated, respondents use the most extreme stimuli to anchor the endpoints of a rating scale. As a result, a given stimulus will be rated as less extreme if presented in the context of a more extreme one, than if presented in the context of a less extreme one. In addition, if the number of stimuli to be rated is large, respondents attempt to use all categories of the rating scale about equally often to be maximally informative. Accordingly, the specific ratings given also depend on the frequency distribution of the presented stimuli. These processes have been conceptualized in a number of related models of rating scale use, of which Parducci's (1965) range-frequency model is the most comprehensive. As a result, ratings of the same object cannot be directly compared when they were collected in different contexts, rendering comparisons over time or between studies difficult.

**Other Scale Format** In representative sample surveys, as well as most psychological experiments, respondents' attitudes toward an object are typically assessed by asking only one or two questions, despite the usual textbook admonition to use multi-item scales. In fact, the classic textbook examples of multi-item attitude scales, like the Thurstone or Guttman scales, are rarely used in practice. All of these scales require extensive topic-specific item development and pretesting to arrive at a set of items that forms an internally consistent scale. Himmelfarb (1993) provides an excellent review of these and other classic scale formats.

In contrast, Osgood and colleagues' (1957) semantic differential scale is a ready-to-use scale that can be applied to any topic without new development work, making it considerably more popular. Respondents are asked to rate the attitude object (e.g., "abortion") on a set of 7-point bipolar adjective scales. The adjectives used as endpoint labels reflect three general factors, namely evaluation (e.g., good-bad; pleasant-unpleasant), potency (e.g., strong-weak; small-large), and activity (e.g., active-passive; fast-slow). Of these factors, evaluation is considered the primary indicator of respondents' attitude toward objects, as reflected in the objects' (relatively global) connotative meanings.

### *Response Editing*

Finally, respondents may hesitate to report their attitude when they are concerned that their answer may present them in a negative light. If so, they may want to edit



their privately formed judgment before they communicate it, essentially providing a more “acceptable” answer. As may be expected, editing on the basis of social desirability is particularly likely when the question is highly threatening (Bradburn et al., 2004; De Maio, 1984). Moreover, it is more pronounced in face-to-face interviews than in self-administered questionnaires, which provide a higher degree of confidentiality (e.g., Krysan et al., 1994; T. W. Smith, 1979).

To reduce socially desirable responding, researchers developed a number of different techniques. Some techniques attempt to ensure the confidentiality of respondents’ answers. Relevant procedures range from simple assurances of anonymity and confidentiality to complex randomized response techniques (Bradburn et al., 2004; Himmelfarb, 1993). In the latter case, respondents are presented with two different questions, an innocuous one and a socially sensitive one, and a draw a card that determines which one they are to answer. Given properly worded response alternatives, the interviewer remains unaware of the question to which the answer pertains, thus ensuring the highest possible level of confidentiality. Other techniques create conditions that present a disincentive for socially desirable responding. For example, Sigall and Page’s (1972) “bogus pipeline” technique involves convincing participants that the researcher can discern their true attitude independent of what they say, thus making lying an embarrassment. Empirically, these various techniques have been found to increase the frequency of socially undesirable answers (Himmelfarb, 1993).

Although socially desirable responding is undoubtedly a threat to the validity of attitude reports, many of the more robust findings commonly attributed to its influence may reflect the impact of several distinct processes. For example, white respondents have frequently been found to mute negative sentiments about African Americans when the interviewer is black rather than white (e.g., Hatchett & Schuman, 1976). From a social desirability perspective, these context dependent answers presumably do not reflect respondents’ “true” attitude. However, the friendly conversation with a middle-class African-American interviewer may itself serve as input into the attitude judgment, resulting in a (temporary) “real” attitude change, much as incidental exposure to pictures or names of liked African Americans has been found to affect attitudes toward the group in laboratory experiments (e.g., Bodenhausen, Schwarz, Bless, & Wänke, 1995). Hence, the impact of social desirability per se often is difficult to isolate. Moreover, social desirability certainly affects everyday behavior, including interracial interactions, indicating that it is not a mere artifact observed in measurement contexts—nor is it obvious that we should disregard social desirability influences when our goal is to predict such everyday behavior.

**Summary** As this selective review indicates, asking people to report on their attitudes will almost always result in an answer—but it often remains unclear what exactly the answer means. Attitude reports are highly context sensitive and minor variations in question wording, format, or sequence can profoundly affect the obtained results. The underlying processes are systematic and increasingly well understood. Sudman and colleagues (1996) and Tourangeau and colleagues (2000) provide comprehensive reviews of what has been learned about the cognitive processes involved in answering questions about one’s attitudes and behaviors.

The observed context dependency of respondents' answers is particularly problematic in survey research. Researchers conduct surveys to generalize from the answers provided by a representative sample to the attitudes of a population that was never exposed to the context in which the sample answered the questions. Hence, any contextual influence on the answers of the sample may lead to erroneous inferences about the population. The problem is less profound in experimental research. In most experiments, we are primarily interested in differences between experimental conditions. As long as the attitude questions (including their format and ordering) are constant across conditions, observed differences between conditions are meaningful, although slightly different questions may have resulted in different answers.

### *Implicit Measures of Attitudes*

Given the context dependency of respondents' answers to direct attitude questions, researchers developed a number of techniques that replace explicit self-reports of attitudes with more indirect measures. The use of indirect measures is based on the theoretical assumption that attitudes exert a systematic influence on people's performance on a variety of tasks and that the size of this influence can serve as an index of the underlying attitude. Accordingly, indirect measures do not require the assumption that people are aware of their attitudes (in contrast to direct questions, which can only be answered on the basis of awareness and introspective insight). To infer a person's attitude from his or her performance on another task, we need clear bridging rules that specify the theoretical and empirical relationship between the attitude and the task. Not surprisingly, these bridging rules have varied widely over the history of attitude research. From the early use of projective tests (e.g., Proshensky, 1943) to the current use of response latency measures (reviewed below), the history of indirect measures mirrors historical shifts in the conceptualization of attitudes and their underlying processes (see Vargas, Sekaquaptewa, & von Hippel, 2007, for an informative review). The respective theoretical assumptions gave rise to numerous controversies, which are beyond the scope of this chapter (see the contributions in Wittenbrink & Schwarz, 2007, for controversies surrounding current reaction time measures).

In addition to requiring no introspective insight into one's attitudes, indirect attitude measures promise to solve the problem of response editing in ways that go beyond what can be achieved in the context of explicit self-reports. First, respondents are presumably unaware of the relationship between their response to indirect measures and their attitudes. Hence, they have few incentives and opportunities for deliberate self-presentation—and wouldn't know how to present themselves in a favorable light even if they wanted to. Second, some researchers have been concerned that deception and self-presentation may not only be directed toward others, but also toward the self (e.g., Paulhus, 1984). From this perspective, people may sometimes hold attitudes of which they are not aware, and hence can't report on, or which they don't even want to admit to themselves. Indirect measures may capture such attitudes because they do not require that respondents are aware

of them; nor does their opaque nature confront the person with the implications of his or her response.

In addition, many researchers also hope that (some) indirect measures may reduce the context dependency observed in explicit attitude reports (for a review see Ferguson & Bargh, 2007). According to one influential conceptualization (Fazio, 1995), attitudes are stored object-evaluation links that are automatically activated upon exposure to the attitude object. From this perspective, context effects reflect noise that results from the deliberate consideration of contextual information, and this noise may be avoided by limiting the degree of deliberate processing (see Ferguson & Bargh, 2007). Hence, fast-paced response latency procedures, which provide little opportunity for deliberation, may limit context effects and may provide a “bona fide pipeline” (Fazio, Jackson, Dunton, & Williams, 1995) to people’s true attitudes.

Next, I turn to these measures and some technologically less demanding paper-and-pencil alternatives. Because these measures do not require awareness of the attitude and entail no explicit attitude report, they are commonly referred to as “implicit” attitude measures (see Devos, this volume, for a more detailed theoretical discussion).

### *Response Time Measures*

The currently most widely used implicit attitude measures rely on response time measurement. Some of these measures take advantage of the observation that preceding exposure to a stimulus facilitates subsequent responses to related stimuli; others draw on the observation that a stimulus is responded to more slowly when it contains multiple features that give rise to competing responses. Bassili (2001) and the contributions in Wittenbrink and Schwarz (2007) provide detailed reviews of these measures and their underlying logic (see Bassili, this volume).

**Sequential Priming Procedures** As a large body of research in cognitive psychology indicates (for a review see Neely, 1991), exposure to a concept (e.g., “doctor”) facilitates the subsequent recognition of related concepts (e.g., “nurse”). A common explanation for this phenomenon holds that exposure to the initial concept (the prime) activates semantically related concepts in memory, thus reducing the time needed for their identification.

Concept priming procedures take advantage of this facilitation effect to assess a person’s associations with an attitude object. They present target words with evaluative meaning (e.g., lazy, smart) and ask participants to identify the word as fast as possible. Speed of identification can be assessed by having participants pronounce the word or by having them decide whether a letter string is a word or a nonword. Of interest is whether a preceding prime that represents the attitude object (e.g., black, white) affects the speed with which different target words can be identified. For example, Wittenbrink, Judd, and Park (1997) exposed participants to African American or white primes and assessed how quickly they could identify subsequently presented trait terms of positive or negative valence that were or were not part of the cultural stereotype about the group. The observed

facilitation patterns provide information that bears on three questions: First, does exposure to the group prime activate associated stereotypical traits, independent of their valence? If so, stereotypical traits will be recognized faster than stereotype unrelated traits. Second, is the automatic activation evaluatively biased? For example, are negative stereotypic traits identified more quickly than positive ones, indicating that the negative traits are more accessible? Third, does exposure to the group prime activate general evaluative associations (e.g., good, bad), independent of their stereotypicality?

*Evaluative priming* procedures (e.g., Fazio, Sanbonmatsu, Powell, & Kardes, 1986) focus on the speed with which the evaluative meaning of a word can be identified. Of interest is whether exposure to the attitude object affects the speed of the evaluative response to the target words. In a typical experiment, participants are exposed to a prime (e.g., a black or white face) and decide whether a subsequent target word (e.g., pleasant, awful) is positive (press the “good” key) or negative (press the “bad” key). If the attitude prime is strongly associated with a positive evaluation, it speeds up the identification of positive words as “good” and slows down the identification of negative words as “bad.” Devos (this volume), Fazio (1995), and Wittenbrink (2007) review representative findings. Unfortunately, the observed facilitation patterns depend to some extent on the general accessibility of the target words and some experiments produced reversals of the usually obtained patterns when the target words have a very high frequency in everyday language use (Chan, Ybarra, & Schwarz, 2006).

In sum, evaluative priming procedures assess whether an attitude object triggers an automatic evaluation, whereas concept priming procedures assess descriptive associations that may have evaluative content. Wittenbrink (2007) reviews these procedures, provides advice on their implementation, and summarizes representative findings.

**Response Competition Procedures** A second class of response time procedures is based on interference effects that may occur when different features of attitude objects imply different responses. The best known of these procedures is the Implicit Associations Test (IAT; Greenwald, McGhee, & Schwartz, 1998). It presents two discrimination tasks that are combined in specific ways across a sequence of five steps. To assess attitudes toward African Americans and European Americans, for example, the first discrimination task may present names that are typical for the respective group and participants are asked to categorize each name as “white” vs. “black.” They do so by pressing a response key assigned to “white” with their left hand or a response key assigned to “black” with their right hand. Next, a second discrimination task presents words with pleasant (e.g., love) or unpleasant (e.g., poison) connotations, which participants classify as positive vs. negative by pressing the left vs. right response key. At the third step, these two tasks are superimposed and participants press the left key when either a white name or a pleasant word is shown, but the right key when either a black name or an unpleasant word is shown. As in the facilitation paradigms, this task is easier when evaluatively associated categories share the same response key; for example, when white participants press the left key to categorize white names and pleasant

words. Going beyond this assessment of response facilitation, the IAT involves two more steps. At the fourth step, the assignment of keys to white and black names is reversed, so that participants who first used the left key for white names now use the left key for black names. Finally, the two discrimination tasks are again superimposed, resulting in an assignment of “black” and “pleasant” to the left response key and “white” and “unpleasant” to the right response key.

Of key interest is the speed with which participants can perform the two superimposed discrimination tasks at step 3 and step 5. Do participants respond faster when a given response key pertains either to the pairing of white names + pleasant words or black names + unpleasant words (step 3) than when this pairing is reversed and a given response key pertains either to white names + unpleasant words or black names + pleasant words (step five)? In the present example, a faster response at step 3 than at step 5 is thought to indicate that white names and positive evaluations and black names and negative evaluations are more strongly associated than the reverse pairings.

Lane, Banaji, Nosek, and Greenwald (2007) review the underlying logic, report representative findings, and provide hands-on advice for the implementation and scoring of the IAT. Related response competition tasks include the Go/No-go Association Task (GNAT; Nosek & Banaji, 2001) and the Extrinsic Affective Simon Task (EAST; De Houwer, 2003).

### *Low Tech Alternatives*

Whereas response time procedures require a high degree of instrumentation and technical sophistication, other implicit measures of attitudes are decidedly low tech. Some of these measures take advantage of the observation that attitudes and expectations influence individuals' information processing in systematic ways. For example, people are more likely to spontaneously explain events that disconfirm rather than confirm their expectations (e.g., Hastie, 1984), suggesting that the amount of explanatory activity can serve as an indirect measure of a person's expectations. The Stereotypic Explanatory Bias (SEB) measure developed by Sekaquaptewa and colleagues (2003) builds on this observation and uses the number of explanations generated in response to stereotype-consistent vs. stereotype-inconsistent behaviors as an implicit measure of stereotyping. Similarly, people describe expected or stereotype-consistent behaviors in more abstract terms than unexpected or stereotype-inconsistent behaviors, a phenomenon known as the Linguistic Intergroup Bias (LIB; e.g., Maass, Salvi, Accuri, & Semin, 1989). The size of this bias can again be used as an indirect measure to gauge the underlying expectations. Vargas and colleagues (2007) review such measures and provide advice on their use.

### *Context Effects on Implicit Measures*

The initial hope that responses to implicit measures that limit deliberation may be less context dependent than responses to explicit attitude questions has not been supported (for reviews see Blair, 2002; Ferguson & Bargh, 2007). Instead, these

measures are subject to pronounced context effects that usually parallel the patterns observed on explicit attitude measures. For example, Dasgupta and Greenwald (2001) observed that exposure to pictures of liked African Americans and disliked European Americans resulted in shifts on a subsequent IAT that parallel the effects of exposure to liked or disliked exemplars on explicit measures of attitudes (e.g., Bodenhausen et al., 1995). Similarly, Wittenbrink, Judd, and Park (2001) found that the same black face primes elicited more negative automatic responses when the faces were presented on the background of an urban street scene rather than a church scene. Other findings parallel the observed influence of interviewer race and ethnicity in the survey research literature (e.g., Hatchet & Schuman, 1976; Weeks & Moore, 1981). For example, Lowery, Hardin, and Sinclair's (2001) obtained more positive automatic evaluations of African Americans when the experimenter was black rather than white. Note that the low transparency of Lowery et al.'s implicit attitude measure makes it unlikely that these responses were based on a deliberate self-presentation strategy. Instead, the accumulating findings suggest that experimenters and interviewers may serve as highly accessible positive exemplars when respondents evaluate the group in general, paralleling the influence of incidental exposure to liked exemplars in other research (e.g., Bodenhausen et al., 1995; Dasgupta & Greenwald, 2001).

To account for the context dependency of implicit measures, Ferguson and Bargh (2007) suggest that automatic attitudes are responses to object-centered contexts rather than to the attitude object in isolation. I return to this issue in the final section of this chapter.

## PSYCHOPHYSIOLOGY AND BEHAVIORAL OBSERVATION

Because of their involuntary and hard to control nature, physiological correlates of evaluative responses have long been of interest to attitude researchers who doubted respondents' explicit self-reports.

### *Psychophysiological Measures*

Early uses of physiological measures drew on the observation that strong affective reactions to an attitude object are associated with increased activation of the sympathetic nervous system (e.g., Rankin & Campbell, 1955). Increased sympathetic activation results in increased sweat glands activity, which can be measured by assessing the resistance of the skin to low level electric currents, a procedure known as electrodermal measurement. However, electrodermal responses do not reflect the direction (favorable or unfavorable) of the evaluative response, which limits their usefulness.

More promising are attempts to assess changes in individuals' facial expression in response to an attitude object. Overt facial expressions (like smiling or frowning) may often be observed in response to attitude objects that elicit strong reactions. But these expressions may be intentionally concealed and many evaluative reactions may be too subtle to evoke overt expressive behaviors. Even subtle evalu-

ative reactions are associated, however, with low-level activation of facial muscles that can be detected by electromyography (EMG). These muscle reactions reflect the direction (favorable vs. unfavorable) as well as the intensity of the evaluative response (see Cacioppo, Bush, & Tassinary, 1992, for an example). However, the obtained measures can be distorted by facial movements that are unrelated to the evaluative reaction.

Another approach involves the measurement of brain activity through electroencephalography (EEG), the assessment of small electric signals recorded from the scalp. This procedure, however, does not lend itself to a direct assessment of positive or negative responses. Instead, it capitalizes on the observation that unexpected stimuli evoke brain wave activity that differs from the activity evoked by expected stimuli. Hence, one may detect if a target object is evaluated positively or negatively by embedding its presentation in a long series of other objects with a known evaluation. The brain activity evoked by the target object will then indicate if its evaluation is consistent or inconsistent with the evaluation of the context objects (see Cacioppo, Crites, Berntson, & Coles, 1993, for an example).

Ito and Cacioppo (2007) review these and other measures, including recent developments in the brain imaging techniques, which provide a promising avenue for future work in this area. Throughout, the implementation of psychophysiological measures requires sophisticated technology and high expertise, and their analysis poses complex issues of data reduction. For these reasons, they are best used in collaboration with a skilled colleague.

### *Behavioral Observation*

In principle, a person's attitude toward some object may be inferred from his or her behavior toward it. However, people's behaviors are influenced by many variables other than their attitudes. Hence, the attitude-behavior relationship is typically weak, unless other variables are taken into account (see Ajzen & Cote, this volume). As a result, the mere observation of a behavior is a poor indicator of the person's attitude per se and behavioral observation is rarely used as a measurement strategy.

## THEORETICAL IMPLICATIONS: CONTEXT DEPENDENCY AND THE NATURE OF ATTITUDES

As this review of different attitude measurement techniques indicates, attitude reports are highly context dependent. This observation holds for traditional "explicit" measures (direct questions) as well as for the more recent "implicit" measures (response time procedures). While the findings are undisputed, their theoretical implications are controversial: Do context effects indicate that attitudes are "constructed" on the spot, based on whatever information is accessible at the time of judgment? Or do they merely reflect some "noise" that does not call the existence of enduring attitudes into question?

Taking the latter position, Eagly and Chaiken (2005, p. 747) suggested that "context effects should be and are pervasive...because attitudinal judgments are

not pure expressions of attitude but outputs that reflect both attitude and the information in the contemporaneous setting.” While the contemporaneous setting gives rise to variability in attitude expression, the “inner state or latent construct that constitutes the attitude can be relatively stable. Therefore, judgments often vary around an average value that is defined by the tendency that constitutes the attitude” (Eagly & Chaiken, 2005, p. 747). In contrast, attitude construction models question the assumption that people “have” enduring attitudes and instead focus on the cognitive and affective processes underlying evaluative judgment (e.g., Lord & Lepper, 1999; Schwarz, 2007; Schwarz & Bohner, 2001; Smith & Conrey, 2007). These models identify the processes that give rise to variation in attitude judgments and specify the conditions under which attitude judgments are stable across time and contexts, as well as the conditions under which attitude judgments predict behavior (see Schwarz, 2007; Schwarz & Bohner, 2001).

More important, both approaches differ in their metatheoretical perspectives (see Schwarz, 2007, for a more detailed discussion). Construal models start with the premise that evaluative judgment stands in the service of action, which requires high context sensitivity. Because action is always located in a specific context, any adaptive system of evaluation should be informed by past experience, but highly sensitive to the specifics of the present. It should overweight recent experience at the expense of more distant experience, and experience from similar situations at the expense of experience from dissimilar situations. In addition, it should take current goals and concerns into account to ensure that the assessment is relevant to what we attempt to do now, in this context. Only such context-sensitive evaluation can guide behavior in adaptive ways by alerting us to problems and opportunities when they exist; by interrupting ongoing processes when needed (but not otherwise); and by rendering information highly accessible that is relevant now, in this situation.

A large body of diverse findings indicates that human cognition meets these requirements (for reviews see Barsalou, 2005; Schwarz, 2002; Smith & Semin, 2004)—and it is presumably no coincidence that the above list of desirable context sensitivities reads like a list of the conditions that give rise to context effects in attitude judgments.

In contrast to this situated construal approach, traditional attitude theories treat attitudes as personal dispositions, that is, “a psychological tendency that is expressed by evaluating a particular entity with some degree of favor or disfavor” (Eagly & Chaiken, 1993, p. 1). In the eyes of critics, this conceptualization derives its considerable intuitive appeal from its compatibility with humans’ pervasive tendency to explain others’ behavior in terms of their dispositions—a tendency otherwise known as the “fundamental attribution error” (Ross, 1977). From a dispositional perspective, context effects are undesirable because they cloud the underlying disposition and undermine the observer’s ability to predict an actor’s behavior on the basis of his or her attitude. Yet from the actor’s own perspective, context sensitive evaluation is an asset, not a liability.

To date, attitude research has predominantly taken the observer’s perspective, deploring the context “dependency” of attitude reports, which presumably obscures the actor’s “true” attitude. Once we adopt the actor’s perspective, deplor-



able context “dependency” turns into laudable context “sensitivity.” If so, there may be more to be learned from exploring the dynamics of context sensitive evaluation than from ever more sophisticated attempts to discover a person’s “true” enduring attitude—attempts that have so far mostly resulted in a reiteration of the same basic lesson: evaluations are context sensitive. Such a shift in theoretical orientation would require a methodological approach to attitude measurement that focuses on evaluation-in-context (Ferguson & Bargh, 2007; Schwarz, 2007), raising new challenges for future research.

## REFERENCES

- Barsalou, L. W. (2005). Situated conceptualization. In H. Cohen & C. Lefebvre (Eds.), *Handbook of categorization in cognitive science* (pp. 619–650). Amsterdam: Elsevier.
- Bassili, J. (2001). Cognitive indices of social information processing. In A. Tesser & N. Schwarz (Eds.), *Blackwell handbook of social psychology: Intraindividual processes* (pp. 68–87). Oxford, UK: Blackwell.
- Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review*, 6, 242–261.
- Bless, H., Schwarz, N., & Wänke, M. (2003). The size of context effects in social judgment. In J. P. Forgas, K. D. Williams, & W. von Hippel (Eds.), *Social judgments: Implicit and explicit processes* (pp. 180–197). Cambridge, UK: Cambridge University Press.
- Bodenhausen, G. V., Schwarz, N., Bless, H., & Wänke, M. (1995). Effects of atypical exemplars on racial beliefs: Enlightened racism or generalized appraisals? *Journal of Experimental Social Psychology*, 31, 48–63.
- Bodenhausen, G. V., & Wyer, R. S. (1987). Social cognition and social reality: Information acquisition and use in the laboratory and the real world. In H. J. Hippler, N. Schwarz, & S. Sudman (Eds.), *Social information processing and survey methodology* (pp. 6–41). New York: Springer Verlag.
- Bradburn, N., Sudman, S., & Wansink, B. (2004). *Asking questions* (2nd ed.). San Francisco: Jossey-Bass.
- Cacioppo, J. T., Bush, L. K., & Tassinary, L. G. (1992). Microexpressive facial reactions as a function of affective stimuli. *Personality and Social Psychology Bulletin*, 18, 515–526.
- Cacioppo, J. T., Crites, S. L., Berntson, G. G., & Coles, M. G. H. (1993). If attitudes affect how stimuli are processed, should they not affect event-related brain potential? *Psychological Science*, 4, 108–112.
- Cantril, H. (1944) *Gauging public opinion*. Princeton, NJ: Princeton University Press.
- Chan, E., Ybarra, O., & Schwarz, N. (2006). Reversing the affective congruency effect: The role of target word frequency of occurrence. *Journal of Experimental Social Psychology*, 42, 365–372.
- Clark, H. H., & Schober, M. F. (1992). Asking questions and influencing answers. In J. M. Tanur (Ed.), *Questions about questions* (pp. 15–48). New York: Russell Sage Foundation.
- Dasgupta, N., & Greenwald, A. G. (2001). On the malleability of automatic attitudes: Combating automatic prejudice with images of liked and disliked individuals. *Journal of Personality and Social Psychology*, 81, 800–814.
- De Houwer, J. (2003). The extrinsic affective Simon task. *Experimental Psychology*, 50, 77–85.

- DeMaio, T. J. (1984). Social desirability and survey measurement: A review. In C. F. Turner & E. Martin (Eds.), *Surveying subjective phenomena* (Vol. 2, pp. 257–281). New York: Russell Sage Foundation.
- Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Fort Worth, TX: Harcourt, Brace, Jovanovich.
- Eagly, A. H., & Chaiken, S. (2005). Attitude research in the 21st century: The current state of knowledge. In D. Albarracin, B. T. Johnson, & M. P. Zanna (Eds.), *The handbook of attitudes* (pp. 743–768). Mahwah, NJ: Erlbaum.
- Fazio, R. (1995). Attitudes as object-evaluation associations: Determinants, consequences, and correlates of attitude accessibility. In R. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 247–282). Mahwah, NJ: Erlbaum.
- Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes. A bona fide pipeline? *Journal of Personality and Social Psychology*, 69, 1013–1027.
- Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, 50, 229–238.
- Ferguson, M. J., & Bargh, J. A. (2007). Beyond the attitude object: Automatic attitudes spring from object-centered-contexts. In B. Wittenbrink & N. Schwarz (Eds.), *Implicit measures of attitudes: Progress and controversies* (pp. 216–246). New York: Guilford.
- Greenwald, A. G., McGhee, D. E., & Schwarz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, 74, 1464–1480.
- Grice, H. P. (1975). Logic and conversation. In P. Cole, & J.L. Morgan (Eds.), *Syntax and semantics: Vol.3. Speech acts* (pp. 41–58). New York: Academic Press.
- Hastie, R. (1984). Causes and effects of causal attribution. *Journal of Personality and Social Psychology*, 46, 44–56.
- Hatchett, S., & Schuman, H. (1976). White respondents and race-of-interviewer effects. *Public Opinion Quarterly*, 39, 523–528.
- Himmelfarb, S. (1993). The measurement of attitudes. In A. H. Eagly & S. Chaiken, *The psychology of attitudes* (pp. 23–43). Fort Worth, TX: Harcourt, Brace, Jovanovich.
- Ito, T. A., & Cacioppo, J. T. (2007). Attitudes as a mental and neural state of readiness: Using physiological measures to study implicit attitudes. In B. Wittenbrink & N. Schwarz (Eds.), *Implicit measures of attitudes: Progress and controversies* (pp. 125–158). New York: Guilford.
- Knäuper, B. (1999). The impact of age and education on response order effects in attitude measurement. *Public Opinion Quarterly*, 63, 347–370.
- Krosnick, J. A., & Alwin, D. F. (1987). An evaluation of a cognitive theory of response order effects in survey measurement. *Public Opinion Quarterly*, 51, 201–19.
- Krosnick, J. A., & Fabrigar, L. R. (in press). *Handbook of questionnaire design*. New York: Oxford University Press.
- Krysan, M., Schuman, H., Scott, L. J., & Beatty, P. (1994). Response rates and response content in mail versus face-to-face surveys. *Public Opinion Quarterly*, 58, 381–399.
- Lane, K. A., Banaji, M. R., Nosek, B. A., & Greenwald, A. G. (2007). Understanding and using the Implicit Association Test: IV. What we know (so far) about the method. In B. Wittenbrink & N. Schwarz (Eds.), *Implicit measures of attitudes: Progress and controversies* (pp. 59–102). New York: Guilford.
- Lord, C. G., & Lepper, M. R. (1999). Attitude representation theory. *Advances in Experimental Social Psychology*, 31, 265–343.
- Lowery, B. S., Hardin, C. D., & Sinclair, S. (2001). Social influence on automatic racial prejudice. *Journal of Personality and Social Psychology*, 81, 842–855.

- Maass, A., Salvi, D., Arcuri, L., & Semin, G. (1989). Language use in intergroup contexts: The linguistic intergroup bias. *Journal of Personality and Social Psychology*, *57*, 981–993.
- Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition*. Hillsdale, NJ: Erlbaum.
- Nosek, B. A., & Banaji, M. R. (2001). The go/no-go association task. *Social Cognition*, *19*, 625–666.
- Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. Urbana, IL: University of Illinois Press.
- Parducci, A. (1965). Category judgments: A range-frequency model. *Psychological Review*, *72*, 407–418.
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, *46*, 598–609.
- Payne, S. L. (1951). *The art of asking questions*. Princeton, NJ: Princeton University Press.
- Proshensky, H. M. (1943). A projective method for the study of attitudes. *Journal of Abnormal and Social Psychology*, *38*, 393–395.
- Rankin, R. E., & Campbell, D. T. (1955). Galvanic skin response to Negro and white experimenters. *Journal of Abnormal and Social Psychology*, *51*, 30–33.
- Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. *Advances in Experimental Social Psychology*, *10*, 173–220.
- Schuman, H., & Presser, S. (1981). *Questions and answers in attitude surveys*. New York: Academic Press.
- Schwarz, N. (1994). Judgment in a social context: Biases, shortcomings, and the logic of conversation. *Advances in Experimental Social Psychology*, *26*, 123–162.
- Schwarz, N. (1996). *Cognition and communication: Judgmental biases, research methods and the logic of conversation*. Hillsdale, NJ: Erlbaum.
- Schwarz, N. (2002). Situated cognition and the wisdom of feelings: Cognitive tuning. In L. Feldman Barrett & P. Salovey (Eds.), *The wisdom in feelings* (pp.144–166). New York: Guilford.
- Schwarz, N. (2007). Attitude construction: Evaluation in context. *Social Cognition*, *25*, 638–656.
- Schwarz, N., & Bless, H. (1992a). Constructing reality and its alternatives: Assimilation and contrast effects in social judgment. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgment* (pp. 217–245). Hillsdale, NJ: Erlbaum.
- Schwarz, N., & Bless, H. (1992b). Scandals and the public's trust in politicians: Assimilation and contrast effects. *Personality and Social Psychology Bulletin*, *18*, 574–579.
- Schwarz, N., & Bless, H. (2007). Mental construal processes: The inclusion/exclusion model. In D. A. Stapel & J. Suls (Eds.), *Assimilation and contrast in social psychology* (pp. 119–141). New York: Psychology Press.
- Schwarz, N. & Bohner, G. (2001). The construction of attitudes. In A. Tesser & N. Schwarz (Eds.), *Blackwell handbook of social psychology: Intraindividual processes* (pp. 436–457). Malden, MA: Blackwell.
- Schwarz, N., & Knäuper, B. (2000). Cognition, aging, and self-reports. In D. Park & N. Schwarz (Eds.), *Cognitive aging. A primer* (pp. 233–252). New York: Psychology Press.
- Schwarz, N., Knäuper, B., Hippler, H. J., Noelle-Neumann, E., & Clark, F. (1991). Rating scales: Numeric values may change the meaning of scale labels. *Public Opinion Quarterly*, *55*, 570–582.
- Schwarz, N., & Oyserman, D. (2001). Asking questions about behavior: Cognition, communication and questionnaire construction. *American Journal of Evaluation*, *22*, 127–160.

- Schwarz, N., Strack, F., & Mai, H.P. (1991). Assimilation and contrast effects in part-whole question sequences: A conversational logic analysis. *Public Opinion Quarterly*, *55*, 3–23.
- Schwarz, N., & Sudman, S. (Eds.) (1996). *Answering questions: Methodology for determining cognitive and communicative processes in survey research*. San Francisco: Jossey-Bass.
- Sekaquaptewa, D., Espinoza, P., Thompson, M., Vargas, P., & von Hippel, W. (2003). Stereotypic explanatory bias: Implicit stereotyping as a predictor of discrimination. *Journal of Experimental Social Psychology*, *39*, 75–82.
- Sigall, H., & Page, R. (1972). Reducing attenuation in the expression of interpersonal affect via the bogus pipeline. *Sociometry*, *35*, 629–642.
- Smith, E. R., & Conroy, F. R. (2007). Mental representations are states not things: Implications for implicit and explicit measurement. In B. Wittenbrink & N. Schwarz (Eds.), *Implicit measures of attitudes: Progress and controversies* (pp. 247–264). New York: Guilford.
- Smith, E. R., & Semin, G. R. (2004). Socially situated cognition: Cognition in its social context. *Advances in Experimental Social Psychology*, *36*, 53–117.
- Smith, T. W. (1979). Happiness. *Social Psychology Quarterly*, *42*, 18–30.
- Strack, F., & Martin, L. (1987). Thinking, judging, and communicating: A process account of context effects in attitude surveys. In H. J. Hippler, N. Schwarz, & S. Sudman (Eds.), *Social information processing and survey methodology* (pp. 123–148). New York: Springer Verlag.
- Strack, F., & Schwarz, N. (2007). Asking questions: Measurement in the social sciences. In M. Ash & T. Sturm (Eds.), *Psychology's territories: Historical and contemporary perspectives from different disciplines* (pp. 225–250). Mahwah, NJ: Erlbaum.
- Sudman, S., Bradburn, N. M., & Schwarz, N. (1996). *Thinking about answers: The application of cognitive processes to survey methodology*. San Francisco, CA: Jossey-Bass.
- Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge, UK: Cambridge University Press.
- Vargas, P. T., Sekaquaptewa, D., & von Hippel, W. (2007). Armed only with paper and pencil: “Low-tech” measures of implicit attitudes. In B. Wittenbrink & N. Schwarz (Eds.), *Implicit measures of attitudes: Progress and controversies* (pp. 103–124). New York: Guilford.
- Weeks, M. F., & Moore, R. P. (1981). Ethnicity-of-interviewer effects on ethnic respondents. *Public Opinion Quarterly*, *45*, 245–249.
- Willis, G. B. (2004). *Cognitive interviewing. A tool for improving questionnaire design*. Thousand Oaks, CA: Sage.
- Wittenbrink, B. (2007). Measuring attitudes through priming. In B. Wittenbrink & N. Schwarz (Eds.), *Implicit measures of attitudes: Progress and controversies* (pp. 17–58). New York: Guilford.
- Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationships with questionnaire measures. *Journal of Personality and Social Psychology*, *72*, 262–274.
- Wittenbrink, B., Judd, C. M., & Park, B. (2001). Spontaneous prejudice in context: Variability in automatically activated attitudes. *Journal of Personality and Social Psychology*, *81*, 815–827.
- Wittenbrink, B., & Schwarz, N. (Eds.). (2007). *Implicit measures of attitudes: Procedures and controversies*. New York: Guilford.
- Wyer, R. S., & Srull, T. K. (1989). *Memory and cognition in its social context*. Hillsdale, NJ: Erlbaum.