6

Survey Response Styles Across Cultures

TIMOTHY P. JOHNSON, SHARON SHAVITT, AND
ALLYSON L. HOLBROOK

INTRODUCTION

Survey reports are susceptible to multiple forms of measurement error (Sudman & Bradburn, 1974; Tourangeau, Rips, & Rasinski, 2000). In this chapter, we consider some of the potential processes through which culture may be implicated in measurement error. In particular, we focus on cultural variability in several common survey response styles, including socially desirable responding (SDR), acquiescent response style (ARS), and extreme response style (ERS). Awareness of response styles is particularly important in the conduct of cross-cultural research. Systematic variance in response style behaviors across racial, ethnic, or national groups may be mistakenly interpreted as cultural differences (or similarities) in the substantive measures being compared (Johnson & van de Vijver, 2003; Keillor, Owens, & Pettijohn, 2001; Middleton & Jones, 2000; Si & Cullen, 1998). Response styles also may suppress or inflate associations among variables (Wells, 1961) differentially across cultural groups. Thus, the potential for cultural variability in survey reporting has direct implications for many academic disciplines that rely on survey research for measurement purposes, as well as for applied researchers working across many substantive fields. This review integrates evidence and experiences from many of these disciplines regarding three of the most common forms of response style that vary across cultures. Three types of evidence are considered: (a) evidence of differences across racial and ethnic groups within nations, (b) evidence of differences across countries, and (c) evidence of associations between direct

measures of cultural values and each response style. We also consider the potential cultural mechanisms underlying these processes. Methodological issues relevant to the measurement of response styles and proposed methods to compensate for cultural heterogeneity in these reporting processes are reviewed as well.

CULTURE AND SOCIALLY DESIRABLE RESPONDING

A widely studied topic in research methodology, SDR continues to be a serious concern in survey measurement because of its potential to introduce response bias (Johnson & van de Vijver, 2003; Paulhus, 1991; Tourangeau & Yan, 2007). SDR is the systematic tendency to give answers that make the respondent look good (Paulhus, 1991). Understanding how social desirability is viewed and pursued in different cultural contexts and groups is key to the validity of cross-cultural research efforts and many other research efforts involving self-reports. In general, research findings indicate that compared with individualists, collectivists have a greater tendency to give responses that make the self look good. This finding has emerged in multiple studies and has been shown across nations, across racial and ethnic groups within nations, and across individual-level cultural variables.

Importantly, recent research suggests that the nature of SDR is multidimensional and that collectivists and individualists may in fact engage in distinct forms of SDR (Lalwani, Shavitt, & Johnson, 2006). That is, there is increasing evidence that SDR comprises the tendency to engage in *impression management*, which is the tendency for people to intentionally misrepresent themselves to appear more favorable, and in *self-deceptive enhancement*, or the tendency for people to have inflated yet genuinely held views of themselves (Paulhus 1998a). Impression management reflects the traditional view of socially desirable responding (Paulhus, 1998a; Schlenker & Britt, 1999; Schlenker, Britt, & Pennington, 1996), a construct often associated with dissimulation or deception (Mick, 1996). We examine the implications of this distinction for cultural differences in SDR later in the chapter.

Definitions of Socially Desirable Responding

As noted, socially desirable responding has traditionally been defined in impression management terms: the reporting by survey respondents of information that projects a favorable image of themselves, sometimes at the expense of accuracy (Nederhof, 1985; Ross & Mirowsky, 1984a). This

reporting style reflects the human propensity to emphasize and occasionally overstate positive qualities and behaviors while deemphasizing or understating negative ones. Survey validation studies generally support this presumption. It has been documented in the United States, for example, that voting (Sigelman, 1982), church attendance (Hadaway, Marler, & Chaves, 1993), and physical exercise (Adams et al., 2005) are overreported, whereas drug use (Fendrich, Johnson, Hubbell, & Spiehler, 2004) and sexually transmitted diseases (Clark, Brasseux, Richmond, Getson, & D'Angelo, 1997) are underreported during survey interviews. Marlowe and Crowne described social desirability in a manner that explicitly conceptualized it as being culturally conditioned, referring to SDR as "the need for social approval and acceptance and the belief that it can be attained by means of *culturally* acceptable and appropriate behaviors" (1961, p. 109) [emphasis added].

Podsakoff, MacKenzie, Lee, and Podsakoff (2003) distinguished between person-centered and survey-item-centered social desirability. They described the former as the tendencies of some individuals to respond to survey items "more as a result of their social acceptability than their true feelings." The latter they defined as "the fact that items may be written in such a way as to reflect more socially desirable attitudes, behavior, or perceptions" (p. 882).

Numerous other investigators have offered definitions of socially desirable responding. Stricker defined SDR as "conformity on items for which the social norms exist and are perceived; conformity on such items simply involves making responses approved by the social norms, i.e., socially desirable responses" (1963, p. 320). Messick defined it as "individual consistencies in the tendency to give answers that are considered desirable by social consensus, a tendency that has been taken by some to mean that the respondent is consciously or unconsciously attempting to place himself or herself in a favorable light" (1991, p. 165).

Measuring Socially Desirable Responding

Measures of social desirability can be placed in two overarching categories. The first type of measure is focused on identifying *individuals'* tendency to engage in socially desirable responding. These measures are based on the assumption that socially desirable responding is a tendency that is systematically greater in some individuals than in others. The second set of measures has focused on identifying social desirable *responding*. These approaches have focused on identifying questions that have social desirability connotations and conditions that foster or discourage socially desirable responding.

Both approaches have been used to identify the extent to which individuals with different characteristics (e.g., race, ethnicity, or cultural orientation) demonstrate socially desirable responding.

Individuals' Socially Desirable Response Tendencies

Several self-report measures have been developed to assess individual differences in socially desirable response tendencies (Crowne & Marlowe, 1964; A. L. Edwards, 1957; Eysenck & Eysenck, 1964; Paulhus, 1998a; Stöber, 2001). These measures have been shown to be associated with reports of favorable and unfavorable opinions and behaviors. For example, persons scoring highly on such measures as the Marlowe–Crowne scale have been found less likely to report unfavorable behaviors and values, such as alcohol consumption, intoxication, and marijuana use (Bradburn & Sudman, 1979), as well as materialism (Mick, 1996).

Some of these measures conceptualize socially desirable responding as a single construct, whereas others conceptualize it as having multiple components. For example, the Paulhus Deception Scales measure both impression management and self-deceptive enhancement (Paulhus 1998a). Impression management is tapped by items such as "I sometimes drive faster than the speed limit" (reverse scored) and "I have never dropped litter on the street" (Paulhus, 1998a). Self-deceptive enhancement is assessed by items such as "My first impressions of people usually turn out to be right" and "I am very confident of my judgments."

In contrast, the Marlowe–Crowne scale measures primarily impression management (Paulhus, 1991). It includes items such as "My table manners at home are as good as when I eat out in a restaurant" and "If I could get into a movie without paying and be sure I was not seen, I would probably do it" (reverse scored). Recent research, however, suggests that the Marlowe–Crowne scale may not tap a single dimension reliably (Leite & Beretvas, 2005).

These scales for tapping SDR have been used to measure differences in the tendency to engage in socially desirable responding across cultural groups (e.g., Johnson & van de Vijver, 2003; Lalwani, Shavitt, & Johnson, 2006) as well as across contexts that heighten the salience of cultural self-views (Lalwani & Shavitt, 2009). However, one difficulty with using these measures in cross-cultural research is that the scales themselves may not be equally applicable across cultural groups. For example, the reliability of the Marlowe–Crowne scale sometimes differs across Western versus Eastern contexts (e.g., internal consistency coefficients of 0.72 vs. 0.43, respectively; Middleton & Jones, 2000).

Most social desirability scales ask respondents to report the extent to which they agree that a series of statements describes them (some for which agreement means greater SDR and others for which agreement means less SDR). An additional approach is to use scenario measures. These measures generally present respondents with a situation (e.g., being offered the opportunity to cheat on a difficult and important class assignment) and then ask them the likelihood that they would engage in a particular response (e.g., choosing to cheat). Such scenario measures have been developed to assess both impression management and self-deceptive enhancement (Lalwani et al., 2006; Lalwani & Shavitt, 2009). One difficulty with using such measures in cross-cultural research is that the measures themselves must be comparable across cultures. Furthermore, the development of scenario measures is difficult and time-consuming, and the final measures may not be useful in populations different from those in which they were developed. For example, a scenario about cheating on a class assignment is reasonable in a study of college students but unlikely to work in a general population study or community sample.

Socially Desirable Response Behavior

Other measures of SDR have focused on measuring socially desirable response behavior. These studies compare answers to questions that do versus those that do not have social desirability connotations. In many studies, questions with social desirability connotations have been identified via face validity (e.g., Aneshensel, Frerichs, Clark, & Yokopenic, 1982; Aquilino, 1998; Groves, 1977; Groves & Kahn 1979; Locander, Sudman, & Bradburn, 1976). A downside to this is that the researcher is determining which questions are likely to be subject to SDR bias within the population being studied. Although this assessment is clear for some questions (e.g., illegal drug use), the social desirability connotations of other questions (e.g., experiencing different types of problems) are more ambiguous and likely to vary across cultures.

One approach to dealing with this difficulty has been to collect empirical data on the social desirability connotations of different questions (e.g., Holbrook, Green, & Krosnick, 2003). Some studies ask respondents which questions they believe others would or would not feel comfortable answering (e.g., Blair, Sudman, Bradburn, & Stocking, 1977; see Johnson & van de Vijver, 2003, for a review). Another approach involves conducting a pretest in which a randomly selected subset of respondents is asked to "fake bad" by giving socially undesirable answers, whereas other respondents are randomly assigned to "fake good" by giving socially desirable answers.

If these two groups give significantly different answers to a question, this indicates that there is a generally agreed-on desirable answer to the question and that it has social desirability connotations (e.g., Holbrook et al., 2003; Wiggins, 1959). Although such approaches allow researchers to identify which questions have social desirability connotations, they require a pretest to do so. Furthermore, cross-cultural research on socially desirable responding would require pretests that establish that the social desirability connotations of questions are consistent across cultures.

Another approach to identifying social desirability connotations has been to examine whether responses to questions are affected by conditions that foster or discourage socially desirable responding. For example, a great deal of evidence indicates that people are more willing to report unfavorable attitudes, beliefs, and behaviors when the reporting circumstances ensure anonymity (Himmelfarb & Lickteig, 1982; Paulhus, 1984; Warner, 1965; see Bradburn, Sudman, & Wansink, 2004, for a review) or greater privacy (e.g., Tourangeau & Smith, 1996; see also Puntoni & Tavassoli, 2007) or when respondents believe researchers have other access to information revealing the truth of their thoughts and actions (e.g., Evans, Hansen, & Mittlemark, 1977; Pavlos, 1972; Sigall & Page, 1971). These approaches have also been used to reduce social desirability response bias (and are discussed later in more detail).

Techniques that increase respondents' anonymity or privacy are useful to identify questions that are likely affected by social desirability connotations. However, they cannot be used to identify specific individuals who are demonstrating socially desirable responding. The characteristics of individuals showing socially desirable responding only can be inferred by comparing the effect of a manipulation (e.g., mode or anonymity) across subgroups of respondents.

Finally, socially desirable responding has been assessed by comparing survey self-reports to "gold standard" validation measures. For example, studies examining reports of voter turnout have compared self-reports of turnout to both official records of whether individual respondents voted (Traugott & Katosh, 1979) and to official reports of turnout in a given election (Clausen, 1968; Traugott & Katosh, 1979). Studies of drug use have compared self-reports of drug use to laboratory tests of drugs found in respondents' hair, saliva, urine, or blood (Colon, Robles, & Sahai, 2001; Fendrich et al., 2004). A variant of this approach involves selecting respondents for a study based on a known characteristic (e.g., having been arrested for driving under the influence) and examining the proportion of respondents who report the behavior (e.g., Locander et al., 1976). This approach allows

researchers to assess whether each individual's response is accurate. However, such gold standards are not available for many of the constructs of interest to researchers (e.g., attitudes, beliefs, and perceptions). Furthermore, in many cases, gold standards such as public records or laboratory tests are not free from error themselves (e.g., errors in turnout records may vary systematically across voting precincts; Traugott, Traugott, & Presser 1992).

Cultural Variables and Socially Desirable Responding

Cross-Cultural Differences in Socially Desirable Responding

There is considerable evidence for differences in SDR as a function of racial and ethnic groups within nations. In the United States, Mexican American and Mexican respondents have been found to provide more socially desirable answers compared with Whites (Ross & Mirowsky, 1984b; Warnecke et al., 1997). Those retaining a strong Mexican identity also are more likely to give socially approved responses, suggesting that social desirability may be mediated by acculturation (Ross & Mirowsky, 1984a). Ross and Mirowsky speculated that the greater tendency to provide socially desirable answers among Mexican Americans (compared with Anglos in the United States) was a consequence of the strong family ties found in Mexican society and related pressures to conform and present a "good face to the outside world" (1984b, p. 190).

U.S.-based research has also documented SDR differences between non-Hispanic Whites and other racial and ethnic populations. Several studies have demonstrated that African American respondents in the United States typically score higher on SDR measures than do non-Hispanic whites (Crandall, Crandall, & Katkovsky, 1965; Fisher, 1967; Klassen, Hornstra, & Anderson, 1975; Warnecke et al., 1997). Similar findings have been reported in South Africa, where Black respondents have been shown to score higher on SDR, compared with White South Africans (D. Edwards & Riordan, 1994).

Validation studies in the United States have also found Black–White differences in the likelihood that self-reported information can be validated using external information sources, such as biological testing and official records. Johnson and Bowman (2003), for example, found that African Americans were consistently less likely to provide accurate information regarding substance use behaviors when compared with non-Hispanic Whites across several dozen studies that they reviewed. African Americans have also been found to underreport abortions (Jones & Forrest, 1992)

and to overreport voting behavior (Abramson & Claggett, 1986; Katosh & Traugott, 1981).

Differences between White and Asian samples in the United States have been reported as well. Non-native Asian college students at an American university have been shown to endorse more Marlowe–Crowne social desirability items than do native White U.S. students (Abe & Zane, 1990). Furthermore, Middleton and Jones (2000) observed more socially desirable answers among college students from East Asian countries studying in the United States compared with U.S.- and Canadian-born students. Keillor et al. (2001) reported higher social desirability scores among Malaysian graduate students compared with students from both the United States and France.

Cross-national studies also have noted SDR differences across nations. Sri Lankan respondents have been shown to provide more socially desirable answers compared with English respondents (Perera & Eysenck, 1984). Malaysian respondents also have been found to express more socially desirable answers than either U.S. or French respondents (no differences in SDR were found between the French and U.S. samples). G. M. Chen (1995) found that Chinese college students disclose less information compared with a sample of American counterparts. Such unwillingness to self-disclose also can reflect an effort to respond desirably (P. B. Smith & Bond, 1998). Finally, in a large multination study, Bernardi (2006) reported lower impression management scores among U.S. college students compared with those from 11 other countries.

However, the nature of the relation between cultural variables and SDR is also dependent on the type of socially desirable responding in question. Lalwani et al. (2006) argued that two distinct response patterns should emerge as a function of cultural orientations or backgrounds – impression management and self-deceptive enhancement (Gur & Sackeim, 1979; Paulhus, 1991; Sackeim & Gur, 1979). Each of these response styles corresponds to different culturally relevant goals. As noted earlier, subscales measuring these dimensions comprise the Paulhus Deception Scales (Paulhus, 1984, 1991, 1998b). Impression management refers to an attempt to present one's self-reported actions in the most positive manner to convey a favorable image (Paulhus, 1998a; Schlenker & Britt, 1999; Schlenker et al., 1996). Concerns about one's social relationships, predominant in collectivistic cultural contexts, likely give rise to a tendency to present oneself in such ways to "fit in" harmoniously and gain social approval.

In contrast, self-deceptive enhancement refers to the tendency to describe oneself in inflated and overconfident terms. It is a predisposition to see one's

skills in a positive light and has been described as a form of "rigid overconfidence" (Paulhus, 1998a). Such a response style is adaptive for individualistic cultural contexts, in which the motive to see oneself as competent and self-reliant predominates because achieving independence is a central goal.

Lalwani et al. (2006) showed that U.S. respondents (individualistic context), compared with those from Singapore (collectivistic context), scored higher in self-deceptive enhancement and lower in impression management. Similarly, European American respondents, compared with Korean American respondents, scored higher in self-deceptive enhancement and lower in impression management. Parallel findings emerged in multicultural student samples when features of the context were manipulated to make salient either an independent self-view (individualistic context) or an interdependent self-view (collectivistic context) (Lalwani & Shavitt, 2009).

Moreover, data in the United States as a function of cultural orientation shed light on the specific cultural goals served by these response styles (Lalwani et al., 2006). Specifically, people with a cultural orientation that emphasizes sociability, benevolence, and cooperation tended to engage in impression management. However, people with a cultural orientation that emphasizes self-competence, self-direction, and independence tended to engage in self-deceptive enhancement. (See Triandis & Gelfand, 1998, and Shavitt, Lalwani, Zhang, & Torelli, 2006, for more on these specific cultural orientations.) The observed response styles thus appear to reflect distinct self-presentational goals – to be seen as sociable and benevolent (impression management) versus self-reliant and capable (self-deceptive enhancement).

Although the evidence for cultural differences in SDR is considerable, some research has failed to find differences in SDR across nations or across ethnic groups within nations. Lai and Linden (1993) reported no differences in social desirability between Asian Canadian and European Canadian respondents on measures of both self-deceptive enhancement and impression management. Heine and Lehman (1995) also reported no differences between samples of European Canadian and Japanese university students in SDR using these same measures. Further, research comparing U.S. and Philippine students failed to identify differences in socially desirable responding (Grimm & Church, 1999). In a nationwide survey of adults in the United States, Gove and Geerken (1977) found no differences in mean social desirability scores between White, African American, and Hispanic respondents. No racial differences in SDR were identified in a probability sample of adults in Erie County, New York, reported by Welte and Russell (1993). Furthermore, Tsushima (1969) reported no differences in

SDR scores between medical patients in New York City of Italian and Irish ethnicity, and Okazaki (2000) found no differences in SDR between White and Asian American college students. Given that several studies have found no SDR differences, research is needed to address more comprehensively the conditions under which cultural and ethnic group differences in SDR would be expected.

Cultural Processes and Socially Desirable Responding

To some extent, the differences in SDR that have emerged can be seen as reflecting cultural differences in the need to maintain harmonious interpersonal relations. This has been identified as an important value within numerous collectivistic cultures, as exemplified by the smooth interpersonal relations style of Filipinos (Church, 1987), the Latino cultural script of *simpatia* (Triandis, Marín, Lisanski, & Betancourt, 1984), and the East Asian courtesy bias (Deutcher, 1973). Indeed, Middleton and Jones (2000) offered several hypotheses linking Hofstede's (2001) cultural constructs with SDR behaviors within the context of Eastern versus Western cultures. They suggested that individuals embedded within high power distance cultures might be more likely to provide socially desirable responses, whereas persons within low power distance environments might feel fewer constraints to self-expression. Uncertainty avoidance, they postulated, might also be associated with SDR, given that individuals in Eastern nations that are characterized by high uncertainty avoidance might be expected to offer socially desirable responses when confronted with ambiguous situations. Persons within feminine cultures might be expected to have higher SDR as well, because those cultures emphasize the maintenance of harmonious social relationships. Members of individualistic societies are expected to feel weaker social pressures to conform and hence be less prone to provide socially desirable answers. In addition, members of societies that emphasize a long-term time orientation, including the long-term maintenance of social relationships, might be expected to feel social pressures to respond in a socially desirable manner.

Bernardi (2006) has offered several hypotheses regarding potential associations between SDR and Hofstede's cultural classifications. He proposed relationships similar to those outlined by Middleton and Jones (2000) with regard to SDR associations with individualism, uncertainty avoidance, and power distance. However, in contrast to Middleton and Jones's expectations, Bernardi proposed that SDR will be greater among more masculine cultures in which competition and the drive for success will produce greater levels of corruption and "an atmosphere of success at any cost" (p. 45). Although

Bernardi reported support at the national level that SDR decreased with sampled countries' individualism scores and increased with uncertainty avoidance, power distance and masculinity were not found to be independently associated with SDR.

Other evidence is available with which to examine some of these hypothesized processes. Van Hemert, van de Vijver, Poortinga, and Georgas (2002) reported significant negative correlations between a nation's individualism score and mean scores on the Lie scale of the Eysenck Personality Inventory (Eysenck & Eysenck, 1964) across 23 nations. In addition, researchers have identified greater propensities among collectivists toward conformity (Bond & Smith, 1996) and unwillingness to self-disclose (P. B. Smith & Bond, 1998), characteristics also likely to be associated with socially desirable reporting (see Johnson & van de Vijver, 2003). Schwartz, Verkasalo, Antonovsky, and Sagiv (1997) reported significant correlations between scores on the Marlowe–Crowne scale and the values of conformity and tradition. In line with this, as noted earlier, collectivistic values have been linked with a tendency to present oneself in normatively appropriate ways (Lalwani et al., 2006).

Another line of research has compared conceptions of what is perceived to be socially desirable (or not) across cultures. High correlations of the perceived social desirability of items between cultural groups have been reported (Diers, 1965; Gough & Heilbrun, 1980; Iwawaki & Cowen, 1964; Williams, Satterwhite, & Saiz, 1998). Strong but variable within-group correlations also have been reported between the perceived desirability of items and their likelihood of being endorsed by respondents within several ethnic and national groups (Dohrenwend, 1966; Iwawaki, Fukuhara, & Hidano, 1966; Phillips & Clancy, 1970; Turk Smith, Smith, & Seymour, 1993). These findings suggest that social desirability may well be a pancultural or etic concept, albeit one that nonetheless reflects varying degrees of cultural conditioning (Johnson & van de Vijver, 2003).

CULTURE AND ACQUIESCENCE

Differences in the propensity to give acquiescent responses have been identified both within racial and ethnic subgroups within nations and also across countries. A great deal of evidence indicates that acquiescent responding is more typical of ethnic minority or non-White respondents than of White Anglo respondents. At the national level, respondents in Asian, Mediterranean, or African societies generally engage in more acquiescent responding than do respondents in North American or Western European countries.

Overall, then, the evidence consistently points to more acquiescent responding among those with collectivistic cultural backgrounds. However, findings for other dimensions of culture, such as uncertainty avoidance and power distance, are somewhat contradictory and variable depending on whether the nation or the respondent is the unit of analysis.

Definitions of Acquiescent Responding

The ARS, also known as "yea-saying," was identified early in the 1900s (Fritz, 1927). Lentz defined it as "the tendency to agree rather than disagree to propositions in general" (1938, p. 659), and Stricker viewed acquiescent response style as one that reflected conformity "on items for which . . . social norms do not exist" (1963, p. 320). In contrast to survey questions for which social norms are more evident and thus invite socially desirable responses, survey items *not* containing obvious social desirability cues, he postulated, may be more likely to elicit acquiescent responses.

Knowles and Condon (1999) outlined two alternative approaches to interpreting acquiescence. According to one approach, acquiescence is a motivational problem of deliberate impression management (Leary & Kowalski, 1990). According to the other approach, acquiescence is the "uncritical acceptance of an item" (Knowles & Condon, 1999, p. 380). Ross and Mirowsky (1984b) agreed with this impression-management interpretation, viewing acquiescence, along with SDR, as an image management technique deliberately used by some respondents in low power positions within a society. However, empirical research has failed to demonstrate an association between ARS and measures of impression management or self-deceptive enhancement (Knowles & Nathan, 1997), although it has been associated with measures of conformity (Heavan, 1983).

The latter view outlined by Knowles and Condon (1999) regarding uncritical acceptance reflects Cronbach's (1942, 1950) earlier conceptualization, which emphasized suboptimal cognitive processing rather than conscious deception. Krosnick's (1991) analysis supports this conceptualization, suggesting that acquiescence is a form of satisficing. In this regard, Knowles and Nathan (1997) have suggested that the cognitive demands that are associated with some survey tasks, such as distractions, audiences, and time pressures, may increase acquiescent responding.

Various elements of the survey itself are thought to be predictive of ARS as well. It has long been believed that survey respondents may be more prone to acquiesce when questions are ambiguous or when respondents are otherwise uncertain of the task (Bass, 1955; Cronbach, 1950; Jackman, 1973;

Moscovici, 1963; Ray, 1983). Stricker (1963) classified ambiguity and read-ability as two question characteristics that may be associated with the ten-dency to provide acquiescent answers. Banta (1961) and Bass (1955) demon-strated empirically that acquiescent responding increased with the ambiguity of survey items. Stricker (1963) also found that acquiescent answers increased for more difficult-to-read attitude, but not personality, questions.

Finally, some have associated acquiescence with demographic charac-teristics of the respondent. For example, ARS has been found to be more common among persons with lower levels of education (Greenleaf, 1992a; Heavan, 1983; McClendon, 1991; Mirowsky & Ross, 1991; Narayan & Kros-nick, 1996; Schuman & Presser, 1981; Watson, 1992).

Measuring Acquiescent Responding

Unlike SDR, there are no multipurpose measures of ARS. A number of approaches have been used in practice to measure this tendency. Most researchers have employed one of the approaches described in this section, but some have combined two or more types of measures into an index of acquiescence response bias (e.g., Baumgartner & Steenkamp, 2001).

Acquiescent responding has been captured most commonly as agreement with heterogeneous sets of survey items, although there is some variation in how "agreement" has been operationalized (e.g., Bachman & O'Malley, 1984). For example, acquiescence has been measured as the percentage of agree–disagree or yes–no questions to which a respondent answered "agree" or "yes" (e.g., Holbrook et al., 2003) or the proportion of statements to which a respondent reported they agreed or strongly agreed (Ross & Mirowsky, 1984b). It also has been measured as the number of "strongly agree" responses minus the number of "strongly disagree" responses to a series of Likert-type questions (Bachman & O'Malley, 1984).

Theoretically, any survey question with positive (e.g., approve, yes, favor, agree) response options and negative (e.g., disapprove, no, oppose, disagree) response options could be subject to acquiescence, but questions using some variant of agree–disagree response options are most common. One difficulty with this agreement-focused approach is that the actual attitude or dimen-sion being measured may contaminate the metric of agreement. In other words, respondents may score high (or low) on such a measure of ARS at least in part because of their attitudes rather than their tendency to acquiesce.

One approach to minimizing contamination of ARS measures has been to examine items on a scale in which some items are reverse coded (e.g., Ross

& Mirowsky, 1984b; Watson, 1992). If the items in a scale all measure a sin-gle underlying construct, respondents should be unlikely to agree with both positively worded items (e.g., those for which greater agreement means being higher on the dimension being measured) and negatively worded items (e.g., those for which greater agreement means being lower on the dimension being measured). Although this approach addresses contami-nation of the ARS measure by the focal dimension being measured, the content of positive and negatively worded items is different.

To address this issue more systematically, researchers have used two approaches to attempt to eliminate contamination. One approach has been to assess ARS by randomly assigning one half of respondents to report whether they agree or disagree with one statement (e.g., "Please tell me whether you agree or disagree with this statement: Individuals are more to blame than social conditions for crime and lawlessness in this country") and the other half to respond to the opposite statement (e.g., "Please tell me whether you agree or disagree with this statement: Social conditions are more to blame than individuals for crime and lawlessness in this country"; Javeline, 1999). Acquiescence is present if the proportion of respondents who agree with one statement is greater than the proportion who disagree with the second. Cross-cultural differences in acquiescent reporting would be assessed by examining whether the effect of question form varies across respondents of different cultural groups or with different cultural values. This method requires that the researcher has control of the content of the questionnaire (i.e., it cannot typically be used to analyze acquiescence in existing data). Furthermore, it is sometimes difficult to ensure that the statements being used are true opposites.

A different approach to measuring ARS has been to compare the responses to an agree–disagree question to responses to a parallel ques-tion using a forced-choice response format (e.g., Javeline, 1999; Narayan & Krosnick, 1996; Schuman & Presser, 1981). For example, Schuman and Presser (1981) analyzed a series of survey experiments in which one half of respondents were randomly assigned an agree–disagree question (e.g., "Please tell me whether you agree or disagree with this statement: Individ-uals are more to blame than social conditions for crime and lawlessness in this country") and the other half were asked a forced-choice question (e.g., "Which in your opinion is more to blame for crime and lawlessness in this country – individuals or social conditions?"; see also Narayan & Kros-nick, 1996). Acquiescence is assessed as the difference in the proportion of respondents who agree with the statement in the agree–disagree format-ted question and the proportion of respondents who choose the parallel

forced-choice option (e.g., "individuals" in our example). As with the previously described method, the challenge with this approach is to make the responses as comparable as possible across the two formats. This assessment approach has not been widely employed for cross-cultural comparisons (although see Javeline, 1999).

Cultural Variables and Acquiescent Responding

Cross-Cultural Differences in Acquiescent Response Style

Within the United States, one of the earliest studies to document ethnic variability in ARS propensity was reported by Lenski and Leggett (1960). In sampling a cross-section of adults in Detroit, they found that 20% of African American respondents gave mutually contradictory answers to survey questions, compared with 5% of White respondents. Subsequent research has confirmed that African Americans generally are more likely to provide acquiescent answers than Whites (Bachman & O'Malley, 1984; Dohrenwend, 1966; Johnson et al., 1997). Latino respondents in the United States have also been found to be more likely to acquiesce (Marín, Gamba, & Marín, 1992; Ross & Mirowsky, 1984b), with the level of acquiescence decreasing with increasing levels of acculturation (Marín et al., 1992). In analyses of a national survey of health care access conducted in the mid-1970s, Aday, Chiu, and Andersen (1980) found higher rates of ARS among Spanish-heritage persons in response to a single pair of positively and negatively worded questions. They reported that 24% of the Spanish-heritage group exhibited "some" tendency to acquiesce, compared with 13% of non-Hispanic Whites and 17% to 18% of other non-White populations. Using a balanced scale to construct a measure of acquiescence, Warnecke et al. (1997) also documented a greater propensity to provide acquiescent responses among African American and Mexican American respondents in the United States, compared with non-Hispanic Whites. Although the trend did suggest elevated levels of ARS among Puerto Ricans relative to non-Hispanic Whites in this study, the difference was not significant. Also, in comparing responses to a questionnaire concerned with beliefs about mental health, Arkoff, Thaver, and Elkind (1966) found that Asian graduate students in Hawaii provided more acquiescent responses than did Whites.

Across countries, evidence is accumulating that acquiescent response styles may be more typical of collectivistic societies (Harzing, 2006; Hofstede, 2001; Johnson, Kulesa, Cho, & Shavitt, 2005; P. B. Smith, 2004; P. B. Smith & Fisher, 2008; van de Vijver, Ploubidis, & van Hemert, 2004). For

instance, across a sample of six European nations, van Herk, Poortinga, and Verhallen (2004) reported higher values on an acquiescence measure among respondents from several Mediterranean countries (Greece, Italy, and Spain), relative to samples of respondents from northwestern European countries (England, France, and Germany). Similar findings were reported by Baumgartner and Steenkamp (2001), who reported that Greek and Portuguese respondents exhibited greater acquiescence than respondents from other European nations. Greeks also were found to have a greater propensity to acquiesce compared with British and Belgian respondents (Steenkamp & Baumgartner, 1998) and U.S. respondents (Triandis & Triandis, 1962). England and Harpaz (1983), in examining large samples of respondents from the United States and Israel, found that differences in response patterns between these two nations also suggested the presence of systematically higher levels of acquiescence among Israeli respondents, which the authors indicated "certainly weakens any inference about meaningful country differences" (p. 55). Further, differences in acquiescent responding have been observed between Kazakh and Russian respondents, a finding interpreted to be a consequence of greater "deference" within the Kazakh culture (Javeline, 1999). In a study reported by Grimm and Church (1999), Philippine college students studying in the United States were found, on average, to acquiesce somewhat more than their U.S. counterparts, but only when the survey instruments were written in English and fewer scale points were provided. South African women also have been reported to acquiesce in their responses more frequently than Canadian adolescent females (Mwamwenda, 1993). Korean college students have also been found to acquiesce more than U.S. college students (Locke and Bail, 2009). In comparing acquiescence scores across 26 nations, analyses by Harzing (2006) found the highest ARS scores among students in Taiwan, Malaysia, India, and Mexico. Taken together, these findings are consistent with the notion of greater ARS among respondents in collectivistic compared with individualistic cultures.

Recent research also provides evidence that the person-level self-construal of independence–interdependence and the cultural-level context of individualism–collectivism may have an interactive effect on ARS. In one of the few studies that investigated the effects of both individual- and cultural-level effects, P. B. Smith and Fisher (2008) reported cross-national and cross-level analyses that demonstrate that interdependent persons are most acquiescent when embedded within collectivistic cultural environments. These findings are useful in specifying more precisely the conditions under which respondents are most likely to respond in an acquiescent manner.

Cultural Processes and Acquiescent Response Style

The findings referenced in the previous section indicate that researchers have begun to identify cultural dimensions that may underlie cross-group variability in ARS. Consistent with this evidence, acquiescence would appear to be more common in collectivistic cultures because they value deference, politeness, and hospitality (Javeline, 1999) and because their belief systems are characterized by holistic and dialectical thinking, which may be more tolerant of contradictory ideas (Choi & Choi, 2002; Minkov, 2008; Watkins & Cheung, 1995; Wong, Rindfleisch, & Burroughs, 2003).

P. B. Smith (2004) has also proposed that acquiescence may be more common within more uncertain and anxiety-prone cultures. Findings linking acquiescence with uncertainty avoidance, however, have been mixed. Smith found that nations scoring high in uncertainty avoidance also scored higher across several measures of acquiescence. He hypothesized that uncertainty-avoidant cultures are more anxiety prone and have less tolerance for ambiguity and that these traits may be associated with this response style. Similar national-level findings have been reported by van de Vijver et al. (2004). Individual-level analyses conducted by Harzing (2006) also suggest a positive association between ARS and uncertainty avoidance. In contrast, hierarchical analyses have found acquiescent responding to be higher within cultures that score low in uncertainty avoidance (Johnson et al., 2005), a finding that supports the view that acquiescence is more common in social environments with greater tolerance for ambiguity and uncertainty.

Findings relevant to potential associations between power distance and ARS are contradictory as well. Multination analyses by Harzing (2006) revealed a positive association between the acquiescent response style and one of three measures of power distance examined in her student samples. These findings were interpreted as suggesting that persons in high power distance cultures would be more likely to defer to persons in positions of authority, an interpretation reinforced by the power disparities in this study between student respondents and faculty investigators. Van de Vijver et al. (2004) reported a similar positive correlation between power distance and acquiescence, but contrary findings were reported by Johnson et al. (2005), who observed a negative relationship between these two constructs in employee samples that were examined via hierarchical models. At the national level, and consistent with other researchers who have examined national-level associations between these measures (Hofstede, 2001; P. B. Smith, 2004; P. B. Smith & Fisher, 2008; van Hemert et al., 2002), they

noted a positive association between Hofstede's measure of power distance and ARS in their sample of nations ($r = 0.21$). However, they found a negative relationship between national-level power distance and individual-level ARS when examined in cross-level models.

Further, Johnson et al. (2005) identified an inverse relationship between cultural masculinity and acquiescent responding. Such findings are consistent with the emphasis placed on assertiveness and decisiveness within masculine cultures (Hofstede, 1998). Other research, however, has failed to find a relationship between this dimension of national culture and ARS at the national level (P. B. Smith, 2004; P. B. Smith & Fisher, 2008; van Hemert et al., 2002).

CULTURE AND EXTREME RESPONSE STYLES

The presence of extreme response propensities has been recognized for many decades (Cronbach, 1946) and is known to vary across racial and ethnic groups and across nations. Within the United States, for example, Latino and African Americans consistently have been shown to exhibit more ERS than do non-Hispanic Whites. Cross-nationally, U.S. samples have been demonstrated to respond in a more extreme manner than East Asian samples, and less so than South Asian samples. Several potential cultural processes are believed to contribute to extreme responding. Among those proposed are differential cultural emphases on sincerity, moderation, modesty, willingness to be judgmental, clarity, assertiveness, and decisiveness during interpersonal communications, as well as familiarity with survey instruments developed within Western scientific traditions.

Definitions of Extreme Responding

Clarke has described ERS as "the tendency for some individuals to consistently use the extreme ends of response scales in a multiple category response format" (2001, p. 302). Similarly, Baumgartner and Steenkamp defined ERS as the "tendency to endorse the most extreme response categories regardless of content" (2001, p. 145). Some earlier psychological research had conceptualized ERS as an indicator of intolerance of ambiguity, suggesting that "a social group with a higher tension level would earn a higher extreme response score than a social group with a lower tension level" (Soueif, 1958, p. 329). Over subsequent decades, however, few researchers have defined ERS in this manner.

Less educated respondents are known to endorse extreme responses more frequently than do those with more education (Greenleaf, 1992a; Marín et al., 1992; Warnecke et al., 1997), as are older respondents (Greenleaf, 1992a; Holbrook et al., 2006). Extreme responding also is believed to be more common when the topics being discussed are more salient to the respondent (Gibbons, Zellner, & Rudek, 1999). Given that the relevance of particular topics is, in many cases, likely to vary across cultures, it is perhaps not surprising that levels of ERS might vary as well.

As with the other response styles discussed in this chapter, cultural variability in ERS propensities may produce differences in empirical findings that are artifactual rather than substantive. Findings from various analytic procedures consequently may become biased. Score frequencies, standard deviations, interitem correlations, and factor structures, for example, may be influenced to differing degrees across groups (Arce-Ferrer, 2006; Arce-Ferrer & Ketterer, 2003; Chun, Campbell, & Yoo, 1974), introducing serious measurement confounds.

Measuring Extreme Response Style

There are several approaches to measuring extreme responding. Most commonly, it is assessed as the proportion of survey items for which a respondent selects an extreme response option. This often is assessed using a measure that was developed for other purposes (Bachman & O'Malley, 1984; Biggs & Das, 1973; Crandall, 1982; Das & Dutta, 1969). However, this approach to measuring ERS only reflects preference for the most extreme response options and does not capture additional variance in respondents' preferences for extremity or moderate response options. Furthermore, using items for which the true response may be correlated may contaminate these measures of ERS (Clarke, 2000; Rutten, De Beuckelaer, & Weijters, 2008). Such ERS indices should be treated in analysis as a binomial proportion, although this usually has not been done (Gold, 1975a, 1975b; Greenleaf, 1992b).

Another approach involves evaluating the variance around mean responses (Greenleaf, 1992a; Hui & Triandis, 1989; Kiesler & Sproull, 1986), again typically using responses to multi-item scales intended for other purposes (e.g., for measuring individual differences). However, this approach also typically involves examining variance across related items (e.g., items in an existing scale that all measure the same underlying construct) and therefore may result in contamination.

Greenleaf (1992b) recommended that existing survey items not be used to measure ERS but that a special set of items be included in a survey to

do so. Accordingly, he developed a measure that seeks to avoid the issue of contamination via a scale in which (a) the items are minimally intercorrelated, (b) the items have similar and known proportions of respondents who chose responses of a given extremity, and (c) "established stochastic techniques and statistical models" can be applied to assess internal reliability or convergent validity (p. 331).

Although it enables measurement of ERS without contamination, the utility of this measure is limited in a number of ways. First and most obviously, it cannot be used to assess ERS in existing data sets that do not include the scale. Second, using this measure necessitates adding a 16-item scale to a survey or study, which is sometimes not financially feasible. In addition, the generalizability of conclusions based on this scale may be limited. For instance, the scale uses 6-point Likert-type response options and "is not intended to measure ERS for scales with different anchors or a different number of response intervals" (Greenleaf, 1992b, p. 330). Furthermore, the items were administered in a mail survey and may not be generalizable to other modes. Thus, it is unclear whether this scale measures a general tendency to select extreme responses across formats or modalities. Researchers using other modes or response formats may need to use the procedures described by Greenleaf (1992b) to develop a new ERS scale. Finally, it is unlikely that all items in the Greenleaf scale are associated with equivalent norms across cultural groups.

Other approaches to measuring ERS have used structural equation modeling (SEM) and item response theory procedures to create a latent factor representing ERS. DeJong, Steenkamp, Fox, and Baumgartner (2008), for example, employed item response theory to develop a measure of ERS that does not assume that all items are equally useful for measuring ERS or require that an item's usefulness for measuring ERS be constant across groups (e.g., countries or racial groups). The authors argued that this approach also allows researchers to use existing survey items to measure ERS even if those items are related (DeJong et al., 2008). This allows researchers to measure ERS without including items (such as those developed by Greenleaf) developed specifically for that purpose.

Cheung and Rensvold (2000) used SEM to assess whether the measurement model for ERS was invariant across cultural groups. Holbrook et al. (2006) argued that although ERS has been defined as a preference for extreme response options regardless of content, measures of ERS used in previous research do not reduce contamination from the content of items used to measure ERS. A measure of ERS was introduced that uses confirmatory factor analysis. With this measure, the extremities of responses to

a series of questions are treated as measured variables that reflect a latent variable representing ERS (with the impact of the latent variable on each measured variable being set to 1.0). Individual differences in ERS were indicated if the variance of this factor was significantly different from zero (Holbrook et al., 2006). In a comparison of this new measure to a measure based on the traditional approaches described earlier, this measure of ERS showed more theoretically sensible associations with demographic and personality variables. Using SEM to measure ERS is a relatively new approach in the literature, but it shows promise for assessing ERS in cross-cultural contexts.

Cultural Variables and Extreme Responding

Cross-Cultural Differences in Extreme Response Style
Several studies have documented variations in extreme responding across racial and ethnic groups within the United States. Early work by Berg and Collier (1953) reported higher levels of ERS among African American versus White male college students. No differences were found, however, for African American versus White females. Using several nationally representative samples of youth participating in the Monitoring the Future project, Bachman and O'Malley (1984) also documented the finding that African American youth were more likely than White youth to select the extreme response categories across several Likert-type scales. Further, Holbrook et al. (2006) reported more extreme responding among African American and Latino samples compared with White non-Hispanic respondents in an analysis of data from the 2004 National Election Survey. Similar findings of higher ERS among African American and Latino than among White college students in the United States were reported by Clarke (2001). In a sample of adults in Chicago, Warnecke et al. (1997) also reported higher ERS among Mexican American and Puerto Rican respondents, relative to non-Hispanic White respondents. Comparisons of patient ratings of medical care have also documented more extreme responding among Hispanic, compared to non-Hispanic, samples (Weech-Maldonado et al., 2008). Additionally, a comparison of Hispanic versus non-Hispanic U.S. Navy recruits reported by Hui and Triandis (1989) indicated higher levels of extreme responding among Hispanics, but only when reporting on 5-point measurement scales; no differences in extreme responding were found when 10-point scales were employed.

ERS differences have also been observed across the few available international comparisons between West and South Asian samples. Stening

and Everett (1984) found more extreme responding among Indonesians, Malaysians, Filipinos, and Thais, relative to U.S. respondents. A seven-nation study conducted by Marshall and Lee (1998) also reported cross-national differences, with samples scoring highest on ERS in the nations of Indonesia and Malaysia and lowest in Australia, the United States, and Singapore (New Zealand and South Korea had intermediate scores). No differences were observed, however, in ERS scores between Philippine and U.S. college students (Grimm & Church, 1999).

In contrast, comparisons between Western and East and Asian samples generally have revealed greater levels of ERS among Western respondents. For example, Chun et al. (1974) identified more ERS within a sample of U.S. college students, compared with a similar sample of Korean students. C. Lee and Green (1991) reported similar findings. C. Chen, Lee, and Stevenson (1995) found increased use of extreme response categories among U.S. students when compared with students from Japan and China. Zax and Takahashi (1967) and Wang et al. (2008) reported similar findings. Among samples of consumers in four nations, those from the United States were found to be more likely to use the full range of 9-point scales, compared with consumers from China, Korea, and Thailand (Yeh et al., 1998). Comparisons between Australian and Asian (predominantly Chinese) samples conducted by Dolnicar and Grun (2007) found higher ERS scores among Australians. Findings consistent with these have also been reported within the United States, where Grandy (1996) reported that Asian American students tended to avoid the extreme endpoints of response scales in comparison to Whites. Gibbons, Hamby, and Dennis (1997) similarly found that U.S.-born college students had higher ERS scores than did a sample of primarily East Asian international students at the same university. J. W. Lee, Jones, Mineyama, and Zhang (2002) found similar results in comparing samples of Caucasian and East Asian (Chinese and Japanese) consumers in Southern California.

Similar findings have been reported in a study by Iwata, Roberts, and Kawakami (1995) that compared samples of Japanese and U.S. adult responses to the Centers for Epidemiologic Studies Depression Scale (CES-D). However, this research noted greater ERS among U.S. respondents for positive items (e.g., symptoms such as feeling good, happy) only. There were no differences in the selection of extreme values when responding to negatively worded depression questions (e.g., feeling fearful or a failure). Iwata and colleagues interpreted these findings as being a consequence of Japan's traditional (and collectivistic) emphasis on the maintenance of social harmony by modestly understating one's personal qualities. Therefore, Iwata

et al.'s research suggests that group differences may be driven more by cultural variability in self-presentation styles than by familiarity with survey question response formats. Consistent with these findings, other research has noted higher levels of nonresponding to positively worded items among Chinese Americans (Ying, 1989).

The pattern of findings reported across the various cultures reviewed in this section are consistent with those obtained from a recent large-scale study by Harzing (2006) that examined response styles across 26 nations. Specifically, she noted more ERS among respondents of the United States, relative to East Asian samples from Taiwan, Japan, Hong Kong, and China. She also observed more extreme responding among South Asians in Malaysia and India and among persons from Mexico and several other Latin American countries, compared with persons in the United States.

Across regions of Europe, Harzing (2006) found higher levels of ERS in southern European and Latin nations, compared with the nations of northern, western, and eastern Europe. These findings also support those from several studies that have included samples from European nations. Among these was one of the earliest comparative investigations of ERS, conducted by Brengelmann (1959), who compared response styles across English and German samples. Results suggested that German respondents systematically provided more extreme responses than did English respondents. Another early cross-national study by Triandis and Triandis (1962) found more extreme responses among Greek students than among U.S. students. A more recent study that examined survey data from six European nations (van Herk et al., 2004) concluded that ERS was more common in the Mediterranean countries of Greece, Italy, and Spain than in the northwestern European nations of France, Germany, and the United Kingdom. A student sample from France, however, engaged in more extreme responding than did one from Australia (Clarke, 2001).

Few studies have examined ERS patterns among African and Middle Eastern populations. Moors (2003) reported that immigrants in Belgium of Moroccan nationality were more likely to select extreme responses than were those with a Turkish background (Moors, 2003). Members of these two minority groups having an oral or reading proficiency (or both) in one of Belgium's languages were more likely to endorse extreme response options. In contrast, a study in Israel by Shapiro, Rosenblood, Berlyne, and Finberg (1976) found greater ERS among Bedouin youth than among their Moroccan Arab peers, who were presumed to have a more Western

orientation. These findings were hypothesized to be a consequence of differential exposure to the English language. These discrepant findings are difficult to reconcile in the absence of additional empirical evidence from similar populations.

Cultural Processes and Extreme Response Style

Several mechanisms have been proposed to underlie the associations between cultural factors and ERS. C. Chen et al. (1995) suggested that a lower propensity to endorse extreme response options in survey questionnaires may be a consequence of East Asians being influenced by the concept of moderation, as emphasized by Confucian philosophy, and by beliefs that individuals should not stand out from their social group. Si and Cullen (1998) have made a similar point, observing that Asian cultures emphasize taking middle positions and avoiding extremes. This is also consistent with East Asian emphases on being interdependent (Markus & Kitayama, 1991), nonjudgmental (Riordan & Vandenberg, 1994), shy (Hoy, 1993), and modest in one's self-presentation (Farh, Dobbins, & Cheng, 1991; Kitayama, Markus, Matsumoto, & Norasakkunit, 1997). It also has been suggested that cultures with a high propensity to respond using extreme response options are those in which sincerity and conviction are emphasized (Arce-Ferrer & Ketterer, 2003; Clarke, 2001; Gibbons et al. 1999) and those with higher tension levels and a corresponding intolerance for ambiguity (Soueif, 1958). East-West cultural differences in dialectical thinking have also been discussed as an explanation for variability in extreme responding (Hamamura, Heine, & Paulhus, 2008; Minkov, 2009).

Observations of acculturation processes may contribute insights to our understanding of how culture influences ERS as well. English–Spanish bilinguals, for example, have been shown to exhibit a greater preference for extreme responses when interviewed in Spanish than in English (Gibbons et al., 1999). Consistent with this, Marín et al. (1992) reported that ERS was inversely correlated with acculturation levels among Latinos in the United States. That is, as Latino immigrants become more acculturated to U.S. society, their propensity to endorse extreme response categories diminished. Marín and colleagues interpreted these findings as being consistent with a fundamental Latino cultural script known as *simpatía* (Triandis et al., 1984), which they argue permits interpersonal interactions "to be more fluid and more responsive to a group's needs when individuals use extreme responses that reflect their unmoderated feelings" (Marín et al., 1992, p. 508). Wang et al. (2008) reported similar trends, finding that U.S.-born

Asian students selected scale endpoints more frequently than did foreign born Asian students.

Arce-Ferrer (2006) further hypothesized that higher ERS propensity among Latinos may be a consequence of an incongruity between collectivist communication patterns that emphasize high-context messages and the more direct individualistic principles underlying survey question construction. Indeed, common survey question design recommendations emphasize the importance of clear, unambiguously worded survey questions (Bradburn et al., 2004) and assume a commonality of direct communication style that is likely to be unwarranted when collecting survey data within high-context cultural groups (Oyserman, Coon, & Kemmelmeier, 2002). Consequently, Arce-Ferrer suggested that among more collectivistic populations, less familiarity with Western-style survey rating scales and the ambiguity inherent in their decontextualized format may result in reduced precision in language use and reduced sophistication in the use of ordinal-style Likert-type response scales. McQuiston, Larson, Parrado, and Flaskerud (2002) also presented evidence suggesting that Likert-type survey question formats may be overly confusing to recently arrived Mexican immigrants. Similar concerns have been expressed by Flaskerud, who, after observing that both Central American and Vietnamese immigrants had difficulty using the response options presented in Likert scales, commented that "the degree of variation Likert-type scales attempt to measure is meaningless to some cultural groups" (1988, p. 186).

Several of Hofstede's (2001) cultural dimensions have been empirically associated with extreme response style as well. An individual-difference measure of intolerance of uncertainty developed by Naemi, Beal and Payne (2009) has, for example, been found to be associated with extreme responding within a sample of U.S. college students. Marshall and Lee (1998) reported a positive correlation between ERS and responses to both a three-item measure of individualism and a six-item measure of religiosity. C. Chen et al. (1995) also found a positive association between a measure of individualism and the likelihood of endorsing extreme response options among students in Japan, Taiwan, and the United States. Other evidence comes from a study by Gibbons et al. (1997), who reported a correlation between ERS and less traditional attitudes toward women's roles, a finding consistent with the suggestion that the use of extreme response options is associated with an individualistic orientation. In a study using scales measuring cultural models of parenting, Lamm and Keller (2007) found more ERS among mothers living in countries with interdependent orientations than

among those in countries with independent or autonomous orientations. This finding was interpreted as being consistent with an understanding of ERS as an expression of "axiomatic, rule-like beliefs that are communally shared" (2007, p. 54). DeJong et al. (2008) also reported a positive association between individualism and ERS in a multilevel analysis conducted with samples representing 26 countries. However, in a more complex finding involving hierarchical models, P. B. Smith and Fisher (2008) reported that *inter*dependent persons were less likely to select extreme responses but that this was especially the case when they were embedded within individualistic cultural environments.

Positive associations with ERS have also been found for national-level measures of uncertainty avoidance (DeJong et al., 2008; Harzing, 2006; van de Vijver et al., 2004), power distance (Johnson et al., 2005), and masculinity (DeJong et al., 2008; Johnson et al., 2005). Taken together, these findings suggest that extreme responding is characteristic of cultures that value distinctive, independent, competitive, assertive, decisive, and sincere behavior and that have a low tolerance for ambiguity (Hamilton, 1968; Marín et al., 1992). Preferences for middling response options – another response style not otherwise discussed in this chapter – may be more common within cultures that value modesty, moderation, interpersonal harmony, and subtlety (Chia, Allred, & Jerzak, 1997). For a review of the middling response style literature, see Yang, Harkness, Chin, and Villar (2010).

STRATEGIES FOR CONFRONTING CULTURAL VARIABILITY IN RESPONSE STYLES

A variety of strategies have been suggested as possible approaches to adjusting for group differences in survey response styles. Three general sets of strategies have been considered. One set emphasizes careful attention to the design of survey items and scales in hopes of eliminating or minimizing the effects of response styles. A second is concerned with addressing elements of the social environment within which survey data are collected that may encourage ARS and SDR. The third set includes several analysis strategies designed to detect or adjust for response style differences across groups. A brief overview of these proposed solutions is provided here. Researchers concerned with cultural variability in response styles would be advised to address this problem when designing measurement instruments and when collecting their data, in addition to employing analytic solutions.

Questionnaire Design Strategies

There are a variety of questionnaire design recommendations intended to minimize cross-group differences in response styles. To minimize socially desirable responding, Jones (1963) suggested avoiding the use of questions that are likely to invite socially desirable responses. T. W. Smith (2003) recommended framing questions to minimize threat and the likelihood that socially desirable responses will be provided, and Mitchell (1973) suggested that "moral" words should not be used when developing survey questions, because they are also likely to encourage socially desirable responding. See also Bradburn et al. (2004) for detailed recommendations.

To avoid ARS, Converse and Presser (1986) recommended that agree–disagree question response formats not be used when constructing survey items. Krosnick (1999) observed that yes–no and true–false formats are vulnerable to ARS as well and recommended using a forced-choice response format to minimize acquiescence. Also commonly suggested is the use of measurement scales that contain balanced sets of positively and negatively worded questions to eliminate or minimize the effects of acquiescence (Cloud & Vaughan, 1970; Jackson, 1967; Javeline, 1999; Knowles & Nathan, 1997; Messick, 1991; Mirowsky & Ross, 1991; Ray, 1979; Watson, 1992). However, Triandis (1972) cautioned that reversed items may have different neutral points across cultures that will further complicate analyses and comparisons. Using an established U.S. consumer scale, Wong et al. (2003) reported that whereas positively worded items and reverse-worded items appear to be largely equivalent in the United States (as assessed by item intercorrelations), they appear to represent different constructs for respondents in Thailand, Japan, Singapore, and Korea. Thus, the authors urged caution in the use of mixed-worded scales cross-culturally.

Other recommendations for minimizing ERS include the use of ranking rather than rating response formats (van Herk, 2000), although respondents may have more difficulty answering ranking items. Research by Diamantopoulos, Reynolds, and Simintiras (2006) also has found differences in ERS to be associated with the types of personal pronouns employed in survey items, suggesting that third-person items may be less susceptible to extreme responding than first-person items. Another suggestion put forth by several researchers is to reduce the length of Likert-type scales to a binary format (Cronbach, 1946, 1950; T. W. Smith, 2003). Interestingly, there is also some evidence that increasing, rather than decreasing, the number of response categories may be the best strategy for minimizing ERS. Experimental work by Hui and Triandis (1989) has suggested that ERS differences across cultural groups can be decreased or eliminated altogether when the number of response options presented to respondents is increased from 5 to 10 options. Similar research by Clarke (2001), however, indicated that cross-national variability in extreme responding actually may increase with the number of response choices available to respondents.

Data Collection Strategies

A rich literature is available suggesting that the race and ethnicity of survey respondents may interact with the race and ethnicity of interviewers to produce acquiescent or socially desirable responses, particularly when questions focus on topics that are racially or ethnically relevant (Anderson, Silver, & Abramson 1988; Schuman & Converse 1971; Stokes-Brown, 2006). Potential explanations include the possibility that respondents defer to the perceived thoughts, feelings, or opinions of interviewers from different cultural backgrounds when answering such questions (Lenski & Leggett, 1960); that respondents wish to portray themselves in a positive or socially desirable manner (Crowne & Marlowe, 1960); because racial stereotypes are activated for the racial or ethnic group of the interviewer (Davis & Silver, 2003); or because of cultural values that dictate politeness to strangers (Jones, 1963). To address these issues, many survey researchers deliberately match respondents with interviewers of a similar racial or ethnic background (Schaeffer, 1980; Vernon, Roberts, & Lee, 1982) or with indigenous interviewers (Bloom & Padilla, 1979) in hopes of placing respondents at greater ease, fostering greater candor, and minimizing the uncertainties of cross-cultural communication (Brislin, 1986). Although that is a common approach, some have argued that employing interviewers from outside may be preferable when conducting research in communities where the need for privacy is strong (Ferketich, Phillips, & Verran, 1993).

Another common strategy for addressing SDR is increasing respondents' privacy and anonymity when answering survey questions; much evidence suggests that social desirability response bias is lessened under conditions of greater privacy or anonymity (e.g., Himmelfarb & Lickteig, 1982; Paulhus, 1984; Tourangeau & Smith, 1996; Warner, 1965). The anonymity and privacy with which respondents are able to answer survey questions have been manipulated in a variety of ways, including the mode in which the survey questions are administered and the use of techniques such as the randomized response technique (RRT; e.g., Warner, 1965) and the item-count technique (ICT; e.g., Droitcour et al., 1991; Miller, Harrel, & Cisin, 1986) that allow respondents to answer a survey question anonymously without an

interviewer or researcher knowing their responses. There is considerable evidence that self-administered surveys result in greater reports of unfavorable attitudes, beliefs, and behaviors and fewer reports of favorable attitudes, beliefs, and behaviors (e.g., Aquilino, 1992; Aquilino & LoSciuto, 1990; Tourangeau & Smith, 1996; see Bradburn et al., 2004, for a review). Similarly, techniques like the RRT and ICT have been shown to reduce socially desirable responding compared with direct self-reports (e.g., Buchman & Tracy, 1982; Dalton, Wimbush, & Daily, 1994; Franklin, 1989; Himmelfarb & Lickteig, 1982). However, applications of these techniques to address SDR concerns in cross-cultural research are rare.

Statistical Adjustment Strategies

More than 40 years ago, D. H. Smith (1967) recognized the importance of addressing the problem of cross-cultural variability in response styles when analyzing survey data. Today, several statistical approaches, many of which were initially developed in monocultural settings, are available as possible strategies for adjusting for cross-group differences in SDR, ARS, and ERS.

Statistical adjustments using standardized measures of social desirability (Middleton & Jones, 2000; Paulhus, 1981, 1991; Pleck, Sonenstein, & Ku, 1996), ARS (Vernon et al., 1982), or ERS (Greenleaf, 1992b) are a simple and commonly employed approach to attempting to compensate for differential response styles across groups. However, some investigators have found this to be an ineffective strategy (Dijkstra, Smit, & Comijs, 2001; Ellingson, Sackett, & Hough, 1999). Latent modeling approaches also have been used. Watson (1992), Billiet and McClendon (2000), and Mirowsky and Ross (1991) each have identified a general ARS dimension via SEM that could be included as a covariate in substantive analyses. Welkenhuysen-Gybels, Billiet, and Cambre (2003) have demonstrated a similar approach to controlling for ARS when assessing measurement equivalence across cultural groups. A related latent-class factor approach has been employed by Moors (2003) to identify and adjust for ERS.

Another popular approach is to apply corrections directly to measures believed to be affected by response style bias. For example, Greenleaf (1992a) described a method that applies a mean correction to item scores to adjust for ARS. Other forms of item and scale standardization have been employed by Leung and Bond (1989), Hofstede (2001), and P. B. Smith (2004). Harzing (2006), however, cautioned that these approaches also may eliminate true substantive differences. Others have attempted to avoid acquiescence and extreme responding effects by developing rescaled ipsative (i.e., ranking)

measures that are believed to be less susceptible to these forms of bias (Cunningham, Cunningham, & Green, 1977; Schuman & Presser, 1981; Toner, 1987). Fischer (2004), Hofstee, ten Berge, and Hendriks (1998), and van de Vijver and Leung (1997) each provided in-depth reviews of these and additional procedures for scoring questionnaires to correct for acquiescence and extreme responding.

Emerging Topics

Despite the rapidly accumulating evidence about the role of culture in survey response styles, much of the evidence can be linked to the broad-based distinction between individualist and collectivist societies. This distinction is profoundly important and thus represents the most broadly used dimension of cultural variability for cross-cultural comparison (Gudykunst & Ting-Toomey, 1988). However, there are limitations to the insights afforded by any broad dimension. Further refinement can stimulate new insights and afford a more nuanced understanding of the link between culture and consumer phenomena (Maheswaran & Shavitt, 2000).

For instance, recent research points to the importance of a relatively new cultural distinction. Triandis and his colleagues (Singelis, Triandis, Bhawuk, & Gelfand, 1995; Triandis, 1995; Triandis, Chen, & Chan, 1998; Triandis & Gelfand, 1998) have proposed that, nested within each individualistic or collectivistic cultural category, some societies are *horizontal* (valuing equality) whereas others are *vertical* (emphasizing hierarchy). This distinction resembles the nation-level power distance construct (Hofstede, 1980, 2001), but there are important conceptual and structural distinctions (see Shavitt et al., 2006).

Although the contribution of the horizontal–vertical distinction is sometimes obscured by methods that conflate it with other dimensions, its impact is distinct from that associated with individualism–collectivism. Across numerous studies, results support the utility of examining distinctions in hierarchy and status motives for the understanding of personal and cultural values, social perceptions, and self-presentational patterns (e.g., Shavitt, Johnson, & Zhang, in press; Torelli & Shavitt, in press; see also Shavitt et al., 2006, for a review). For instance, people with a horizontal collectivistic cultural orientation, who emphasize sociability, benevolence, and cooperation, are characterized by a tendency to engage in impression management across multiple survey measures of this tendency. However, people with a vertical collectivistic orientation, who emphasize stature, duty, and conformity, are not especially likely to be concerned with impression management

(Lalwani et al., 2006). One might speculate that the vertical collectivistic orientation would instead be more predictive of deferential responding, perhaps including acquiescent responding, particularly if a survey is fielded by a person or organization toward which the respondent feels a sense of obligation or duty.

People with a horizontal individualistic orientation, who emphasize self-competence, self-direction, and independence, have a tendency to engage in self-deceptive enhancement (SDE). In contrast, those with a vertical individualistic orientation, who put emphasis on status, achievement, and personal power (Torelli & Shavitt, in press) are not likely to exhibit SDE (Lalwani et al., 2006). Instead, one may speculate that the vertical individualistic orientation would be more predictive of desirable self-presentations concerning one's achievements and competitive success, as well as a greater tendency to employ extreme response categories and a lesser inclination to provide acquiescent answers.

Recent research on a closely related construct, *status differentiation* (Matsumoto, 2007), suggests that this dimension also accounts for unique variance in cross-national differences beyond the role of other values such as collectivism. In this research, a 20-item scale was developed to address how individuals differentiate their self-regulation and their assertive behaviors toward others depending on the status inequalities between them. The results point to cultural distinctions in how differently people claim to treat those above them versus those below them in status, with Japanese showing a more hierarchical form of status differentiation than U.S. and South Korean participants. Therefore, status differentiation may be associated with a tendency to acquiesce in response to interviewers or survey organizations that one perceives to be above oneself in status. It may also be associated with assertive and extreme responding in response to interviewers or survey organizations that one perceives to be below oneself in status.

In sum, the constructs of horizontal and vertical individualism and collectivism, and the related dimension of status differentiation, hold promise for illuminating how power and status motivations influence survey response styles. The impact of these variables on survey response styles across cultures is a topic worthy of future research.

DISCUSSION

This chapter has reviewed a growing body of evidence suggesting that culture is an important mediator of several commonly recognized survey response styles, including acquiescent, extreme, and socially desirable responding.

This research, distilled from a diverse range of disciplines, has produced generally consistent patterns of findings, which are all the more impressive given the multiple levels and diverse types of cultural indicators that have been examined. For example, it would appear that SDR, broadly defined, is more common both among less powerful cultural groups within societies and among more traditional (i.e., collectivistic) cultures across nations. Similarly, acquiescent responding is more common both among cultural minority groups within nations and among respondents from more collectivistic and uncertainty-avoidant nations. Systematic differences in extreme responding also appear between Western and East Asian cultures. These may represent different degrees of individualism and uncertainty avoidance. (ERS differences between White and minority populations in the United States, however, are less well understood.)

The largely consistent nature of these findings is noteworthy given the multiple measures of each of these response styles that have been employed in practice. However, some of the cultural variability in response styles identified so far may be confounded with culture-based differences in how survey questions are interpreted or in substantive differences across cultures. A general limitation, then, is that most of the cross-cultural comparisons of response style measures reviewed have not investigated the role of measurement equivalence. Although this is a general concern, it would appear to be particularly problematic when comparing measures of social desirability across cultures. It will be important for future research to address this oversight.

It is also important to recognize that nearly all research concerned with cultural variability in survey response styles has been opportunistic in nature. We are aware of few studies that were designed explicitly to investigate response style variability cross-culturally (c.f., Lalwani et al., 2006). Rather, most research appears to have comprised creative secondary analyses of substantive survey data that were primarily collected for other purposes. Although these analyses have been highly informative, it may be difficult to make further advances without specifically designing research to confront these issues. Perhaps of primary importance would be research that clearly documents how failure to recognize or properly adjust for cross-cultural differences in survey response styles contributes erroneous inferences. Although such problems undoubtedly exist in the literature, the research reviewed here is only able to document cultural differences successfully using scales designed specifically to measure one or more response styles. Examples in which these differences actually altered interpretations of substantive measures are relatively rare (although see Iwata et al., 1995).

Finally, it should be acknowledged that several other survey-related response styles exist that were not reviewed in this chapter. Some of these include a middling response style (Chia et al., 1997; Si & Cullen, 1998), disacquiescence (Baumgartner & Steenkamp, 2001), nondifferentiation (Holbrook et al., 2003) and random, or noncontingent, responding (Watkins & Cheung, 1995). To date, these have received comparatively little attention; thus, little is known about the role of culture in these survey response patterns.

REFERENCES

Abe, J. S., & Zane, N. W. S. (1990). Psychological maladjustment among Asian and White American college students: Controlling for confounds. *Journal of Counseling Psychology, 37,* 437–444.

Abramson, P. R., & Claggett, W. (1986). Race-related differences in self-reported and validated turnout in 1984. *Journal of Politics, 48,* 412–422.

Adams, S. A., Matthews, C. E., Ebbeling, C. B., Moore, C. G., Cunningham, J. E., Fulton, J., et al. (2005). The effect of social desirability and social approval on self-reports of physical activity. *American Journal of Epidemiology, 161,* 389–398.

Aday, L. A., Chiu, G. Y., & Andersen, R. (1980). Methodological issues in health care surveys of the Spanish heritage population. *American Journal of Public Health, 70,* 367–374.

Anderson, B. A., Silver, B. D., & Abramson, P. R. (1988). The effects of race of the interviewer on measures of electoral participation by blacks in SRC National Election Studies. *Public Opinion Quarterly, 52,* 53–83.

Aneshensel, C. S., Frerichs, R. R., Clark, V. A., & Yokopenic, P. A. (1982). Measuring depression in the community: A comparison of telephone and personal interviews. *Public Opinion Quarterly, 46,* 110–121.

Aquilino, W. S. (1992). Telephone versus face-to-face interviewing for household drug use surveys. *International Journal of Addiction, 27,* 71–91.

Aquilino, W. S. (1998). Effects of interview mode on measuring depression in younger adults. *Journal of Official Statistics, 14,* 15–29.

Aquilino, W. S., & LoSciuto, L. A. (1990). Effects of interview mode of self-reported drug use. *Public Opinion Quarterly, 54,* 362–395.

Arce-Ferrer, A. J. (2006). An investigation into the factors influencing extreme-response style: Improving meaning of translated and culturally adapted rating scales. *Educational and Psychological Measurement, 66,* 374–392.

Arce-Ferrer, A. J., & Ketterer, J. J. (2003). The effect of scale tailoring for cross-cultural application on scale reliability and construct validity. *Educational and Psychological Measurement, 63,* 484–501.

Arkoff, A., Thaver, F., & Elkind, L. (1966). Mental health and counseling ideas of Asian and American students. *Journal of Counseling Psychology, 13,* 219–223.

Bachman, J. G., & O'Malley, P. M. (1984). Yea-saying, nay-saying, and going to extremes: Black–White differences in response style. *Public Opinion Quarterly, 48,* 491–509.

Banta, T. J. (1961). Social attitudes and response styles. *Educational and Psychological Measurement, 21,* 543–557.

Bass, B. M. (1955). Authoritarianism or acquiescence? *Journal of Abnormal and Social Psychology, 51,* 616–623.

Baumgartner, H., & Steenkamp, J.-B. E. M. (2001). Response styles in marketing research: A cross-national investigation. *Journal of Marketing Research, 38,* 143–156.

Berg, I. A., & Collier, J. S. (1953). Personality and group differences in extreme response sets. *Educational and Psychological Measurement, 13,* 164–169.

Bernardi, R. A. (2006). Associations between Hofstede's cultural constructs and social desirability response bias. *Journal of Business Ethics, 65,* 43–53.

Biggs, J. B., & Das, J. P. (1973). Extreme response set, internality–externality, and performance. *British Journal of Social and Clinical Psychology, 12,* 99–210.

Billiet, J. B., & McClendon, M. J. (2000). Modeling acquiescence in measurement models for two balanced sets of items. *Structural Equation Modeling, 7,* 608–628.

Blair, E., Sudman, S., Bradburn, N. M., & Stocking, C. (1977). How to ask questions about drinking and sex: Response effects in measuring consumer behavior. *Journal of Marketing Research, 14,* 316–321.

Bloom, D., & Padilla, A. M. (1979). A peer interviewer model in conducting surveys among Mexican-American youth. *Journal of Community Psychology, 7,* 129–136.

Bond, R., & Smith, P. B. (1996). Culture and conformity: A meta-analysis of studies using Asch's (1952b, 1956) line judgment task. *Psychological Bulletin, 119,* 111–137.

Bradburn, N. M., & Sudman, S. (1979). Reinterpreting the Marlowe–Crowne scale. In N. M. Bradburn & S. Sudman (Eds.), *Improving interview method and questionnaire design* (pp. 85–106). San Francisco: Jossey-Bass.

Bradburn, N., Sudman, S., & Wansink, B. (2004). *Asking questions: The definitive guide to questionnaire design – For market research, political polls, and social and health questionnaires* (rev ed.). San Francisco: Jossey-Bass.

Brengelmann, J.C. (1959). Differences in questionnaire responses between English and German nationals. *Acta Psychologica, 16,* 339–355.

Brislin, R. W. (1986). The wording and translation of research instruments. In W. J. Lonner & J. W. Berry (Eds.), *Field methods in cross-cultural research* (pp. 137–164). Beverly Hills, CA: Sage.

Buchman, T. A., & Tracy, J. A. (1982). Obtaining responses to sensitive questions: Conventional questionnaire versus randomized response technique. *Journal of Accounting Research, 20,* 263–271.

Chen, C., Lee, S.-Y., & Stevenson, H. W. (1995). Response style and cross-cultural comparisons of rating scales among East Asian and North American students. *Psychological Science, 6,* 170–175.

Chen, G.-M. (1995). Differences in self-disclosure patterns among Americans versus Chinese: A comparative study. *Journal of Cross-Cultural Psychology, 26,* 84–91.

Cheung, G. W., & Rensvold, R. B. (2000). Assessing extreme and acquiescence response sets in cross-cultural research using structural equations modeling. *Journal of Cross-Cultural Psychology, 31,* 187–212.

Chia, R. C., Allred, L. J., & Jerzak, P. A. (1997). Attitudes toward women in Taiwan and China. *Psychology of Women Quarterly, 21*, 137–150.

Chun, K.-T., Campbell, J. B., & Yoo, J. H. (1974). Extreme response style in cross-cultural research: A reminder. *Journal of Cross-Cultural Psychology, 5*, 465–480.

Church, A. T. (1987). Personality research in a non-Western culture: The Philippines. *Psychological Bulletin, 102*, 272–292.

Clark, L. R., Brasseux, C., Richmond, D., Getson, P., & D'Angelo, L. J. (1997). Are adolescents accurate in self-report of frequencies of sexually transmitted diseases and pregnancies? *Journal of Adolescent Health, 21*, 91–96.

Clarke, I., III. (2000). Extreme response style in cross-cultural research: An empirical investigation. *Journal of Social Behavior & Personality, 15*, 137–152.

Clarke, I., III. (2001). Extreme response style in cross-cultural research. *International Marketing Review, 18*, 301–324.

Clausen, A. R. (1968). Response validity – Vote report. *Public Opinion Quarterly, 32*, 588–606.

Cloud, J., & Vaughan, G. M. (1970). Using balanced scales to control acquiescence. *Sociometry, 33*, 193–202.

Colon, H. M., Robles, R. R., & Sahai, H. (2001). The validity of drug use responses in a household survey in Puerto Rico: Comparison of survey responses of cocaine and heroin use with hair tests. *International Journal of Epidemiology, 30*, 1042–1049.

Converse, J. M., & Presser, S. (1986). *Survey questions: Handcrafting the standardized questionnaire.* Thousand Oaks, CA: Sage.

Crandall, J. E. (1982). Social interest, extreme response style, and implications for adjustments. *Journal of Research in Personality, 16*, 82–89.

Crandall, V. C., Crandall, V. J., & Katkovsky, W. (1965). A children's social desirability questionnaire. *Journal of Consulting Psychology, 29*, 27–36.

Cronbach, L. J. (1942). Studies of acquiescence as a factor in the true–false test. *Journal of Educational Psychology, 33*, 401–415.

Cronbach, L. J. (1946). Response sets and test validity. *Educational and Psychological Measurement, 6*, 475–494.

Cronbach, L. J. (1950). Further evidence on response sets and test design. *Educational and Psychological Measurement, 10*, 3–31.

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of pathology. *Journal of Consulting Psychology, 24*, 349–354.

Cunningham, W., Cunningham, I. C., & Green, R. T. (1977). The ipsative process to reduce response set bias. *Public Opinion Quarterly, 41*, 379–394.

Dalton, D. R., Wimbush, J. C., & Daily, C. M. (1994). Using the unmatched count technique (UCT) to estimate base rates for sensitive behavior. *Personnel Psychology, 47*, 817–828.

Das, J. P., & Dutta, T. (1969). Some correlates of extreme response set. *Acta Psychologica, 29*, 85–92.

Davis, D. W., & Silver, B. D. (2003). Stereotype threat and race of interviewer effects in a survey on political knowledge. *American Journal of Political Science, 47*, 33–45.

DeJong, M. G., Steenkamp, J.-B. E. M., Fox, J. P., & Baumgartner, H. (2008). Using item response theory to measure extreme response style in marketing research: A global investigation. *Journal of Marketing Research, 45*, 104–115.

Deutscher, I. (1973). Asking questions: Linguistic comparability. In D. Warwick & S. Osherson (Eds.), *Comparative research methods* (pp. 163–186). Englewood Cliffs, NJ: Prentice-Hall.

Diamantopoulos, A., Reynolds, N. L., & Simintiras, A. C. (2006). The impact of response styles on the stability of cross-national comparisons. *Journal of Business Research, 59*, 925–935.

Diers, C. J. (1965). Social-desirability ratings of personality items by three subcultural groups. *Journal of Social Psychology, 67*, 97–104.

Dijkstra, W., Smit, J. H., & Comijs, H. C. (2001). Using social desirability scales in research among the elderly. *Quality & Quantity, 35*, 107–115.

Dohrenwend, B. (1966). Social status and psychological disorder: An issue of substance and an issue of method. *American Sociological Review, 31*, 14–34.

Dolnicar, S., & Grun, B. (2007). Cross-cultural differences in survey response patterns. *International Marketing Review, 24*, 127–143.

Droitcour, J., Caspar, R. A., Hubbard, M. L., Parsley, T. L., Visscher, W., & Ezzati, T. M. (1991). The item count technique as a method of indirect questioning: A review of its development and a case study application. In P. B. Biemer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz, & S. Sudman (Eds.), *Measurement errors in surveys* (pp. 185–210). New York: Wiley.

Edwards, A. L. (1957). *The social desirability variable in personality assessment and research.* New York: Druiden Press.

Edwards, D., & Riordan, S. (1994). Learned resourcefulness in Black and White South African university students. *Journal of Social Psychology, 134*, 665–675.

Ellingson, J. E., Sackett, P. R., & Hough, L. M. (1999). Social desirability corrections in personality measurement: Issues of applicant comparison and construct validity. *Journal of Applied Psychology, 84*, 155–166.

England, G. W., & Harpaz, I. (1983). Some methodological and analytic considerations in cross-national comparative research. *Journal of International Business Studies, 14*, 49–59.

Evans, R. I., Hansen, W. B., & Mittlemark, M. B. (1977). Increasing the validity of self-reports of smoking behavior in children. *Journal of Applied Psychology, 62*, 521–523.

Eysenck, H. J., & Eysenck, S. B. G. (1964). *The manual of the Eysenck Personality Inventory.* London: University of London Press.

Farh, J., Dobbins, G. H., & Cheng, B. (1991). Cultural relativity in action: A comparison of self-ratings made by Chinese and U.S. workers. *Personnel Psychology, 44*, 129–147.

Fendrich, M., Johnson, T. P., Hubbell, A., & Spiehler, V. (2004). The utility of drug testing in epidemiological research: Results from an ACASI general population survey. *Addiction, 99*, 197–208.

Ferketich, S., Phillips, L., & Verran, J. (1993). Development and administration of a survey instrument for cross-cultural research. *Research in Nursing & Health, 16*, 227–230.

Fischer, R. (2004). Standardization to account for cross-cultural response bias: A classification of score adjustment procedures and review of research in JCCP. *Journal of Cross-Cultural Psychology, 35*, 263–282.

Fisher, G. (1967). The performance of male prisoners on the Marlowe–Crowne social desirability scale: II. Differences as a function of race and crime. *Journal of Clinical Psychology, 23*, 473–475.

Flaskerud, J. (1988). Is the Likert scale format culturally biased? *Nursing Research, 37*, 185–186.

Franklin, L. A. (1989). Randomized response sampling from dichotomous populations with continuous randomization. *Survey Methodology, 15*, 225–235.

Fritz, M. F. (1927). Guessing in a true-false test. *Journal of Educational Psychology, 18*, 558–561.

Gibbons, J. L., Hamby, B. A., & Dennis, W. D. (1997). Researching gender-role ideologies internationally and cross-culturally. *Psychology of Women Quarterly, 21*, 151–170.

Gibbons, J. L., Zellner, J. A., & Rudek, D. J. (1999). Effects of language and meaningfulness on the use of extreme response style by Spanish-English bilinguals. *Cross-Cultural Research: The Journal of Comparative Social Science, 33*, 369–381.

Gold, B. (1975a). A comment on Shulman's "Comparison of two scales on extremity response bias." *Public Opinion Quarterly, 39*, 123–124.

Gold, B. (1975b). Reply to Art Shulman. *Public Opinion Quarterly, 39*, 126–127.

Gough, H. G., & Heilbrun, A. B. (1980). *The adjective check list manual.* Palo Alto, CA: Consulting Psychologists Press.

Gove, W. R., & Geerken, M. R. (1977). Response bias in surveys of mental health: An empirical investigation. *American Journal of Sociology, 82*, 1289–1317.

Grandy, J. (1996). *Differences in the survey responses of Asian American and White science and engineering students* (GRE Board Report No. 93-25P). Princeton, NJ: Educational Testing Service.

Greenleaf, E. A. (1992a). Improving rating scale measures by detecting and correcting bias components in some response styles. *Journal of Marketing Research, 29*, 176–188.

Greenleaf, E. A. (1992b). Measuring extreme response style. *Public Opinion Quarterly, 56*, 328–351.

Grimm, S. D., & Church, A. (1999). A cross-cultural study of response biases in personality measures. *Journal of Research in Personality, 33*, 415–441.

Groves, R. M. (1977). An experimental comparison of national telephone and personal interview surveys. In *Proceedings of the Section on Social Statistics: American Statistical Association* (pp. 232–241). Washington, DC: American Statistical Association.

Groves, R. M., & Kahn, R. L. (1979). *Surveys by telephone: A national comparison with personal interviews.* New York: Academic Press.

Gudykunst, W. B., & Ting-Toomey, S. (1988). *Culture and interpersonal communication.* Newbury Park, CA: Sage.

Gur, R. C., & Sackeim, H. A. (1979). Self-deception: A concept in search of a phenomenon. *Journal of Personality and Social Psychology, 37*, 147–169.

Hadaway, K., Marler, P., & Chaves, M. (1993). What the polls don't show: A closer look at U.S. church attendance. *American Sociological Review, 58*, 741–752.

Hamilton, D. L. (1968). Personality attributes associated with extreme response style. *Psychological Bulletin, 69*, 192–203.

Hamamura, T., Heine, S.J., & Paulhus, D.L. (2008). Cultural differences in response styles: The role of dialectical thinking. *Personality and Individual Differences, 44*, 932–942.

Harzing, A. W. (2006). Response styles in cross-national survey research: A 26-country study. *International Journal of Cross-Cultural Management, 6*, 243–266.

Heavan, P. C. L. (1983). Authoritarianism or acquiescence? South African findings. *Journal of Social Psychology, 119*, 11–15.

Heine, S. J., & Lehman, D. R. (1995). Social desirability among Canadian and Japanese students. *Journal of Social Psychology, 135*, 777–779.

Himmelfarb, S., & Lickteig, C. (1982). Social desirability and the randomized response technique. *Journal of Personality and Social Psychology, 43*, 710–717.

Hofstede, G. H. (1980). *Culture's consequences: International differences in work-related values.* Newbury Park: Sage.

Hofstede, G. H. (1998). *Masculinity and femininity: The taboo dimension of national cultures.* Thousand Oaks, CA: Sage.

Hofstede, G. H. (2001). *Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations.* Thousand Oaks, CA: Sage.

Hofstee, W., ten Berge, J., & Hendriks, A. (1998). How to score questionnaires. *Personality and Individual Differences, 25*, 897–909.

Holbrook, A. L., Green, M. C., & Krosnick, J. A. (2003). Telephone vs. face-to-face interviewing of national probability samples with long questionnaires: Comparisons of respondent satisficing and social desirability response bias. *Public Opinion Quarterly, 67*, 79–125.

Holbrook, A. L., Johnson, T. P., & Cho, Y. I. (2006, May). *Extreme response style: Style or substance?* Paper presented at the 61st Annual Meeting of the American Association for Public Opinion Research, Montreal, Canada.

Hoy, R.R. (1993). A 'model minority' speaks out on cultural shyness. *Science, 262*, 1117–1118.

Hui, C., & Triandis, H. C. (1989). Effects of culture and response format on extreme response style. *Journal of Cross-Cultural Psychology, 20*, 296–309.

Iwata, N., Roberts, C. R., & Kawakami, N. (1995). Japan–U.S. comparison of responses to depression scale items among adult workers. *Psychiatry Research, 58*, 237–245.

Iwawaki, S., & Cowen, E. L. (1964). The social desirability of trait-descriptive terms: Applications to a Japanese sample. *Journal of Social Psychology, 63*, 199–205.

Iwawaki, S., Fukuhara, M., & Hidano, T. (1966). Probability of endorsement of items in the Yatabe–Guilford Personality Inventory: Replication. *Psychological Reports, 19*, 249–250.

Jackson, D. N. (1967). Acquiescent response styles: Problems of identification and control. In I. A. Berg (Ed.), *Response set in personality assessment* (pp. 71–114). Chicago: Aldine.

Jackman, M. R. (1973). Education and prejudice or education and response set? *American Sociological Review, 38*, 327–339.

Javeline, D. (1999). Response effects in polite cultures. *Public Opinion Quarterly, 63*, 1–28.

Johnson, T. P., & Bowman, P. J. (2003). Cross-cultural sources of measurement error in substance use surveys. *Substance Use & Misuse, 38,* 1447–1490.

Johnson, T. P., Kulesa, P., Cho, Y. I., & Shavitt, S. (2005). The relation between culture and response styles: Evidence from 19 countries. *Journal of Cross-Cultural Psychology, 36,* 264–277.

Johnson, T. P., O'Rourke, D., Chávez, N., Sudman, S., Warnecke, R., Lacey, L., et al. (1997). Social cognition and responses to survey questions among culturally diverse populations. In L. Lyberg, P. Biemer, M. Collins, E. de Leeuw, C. Dippo, N. Schwarz, & D. Trewin (Eds.), *Survey measurement and process quality* (pp. 87–113). New York: Wiley.

Johnson, T. P., & Van de Vijver, F. J. (2003). Social desirability in cross-cultural research. In J. A. Harkness, F. J. Van de Vijver, & P. P. Mohler (Eds.), *Cross-cultural survey methods* (pp. 195–204). New York: Wiley.

Jones, E. L. (1963). The courtesy bias in South-East Asian surveys. *International Social Science Journal, 15,* 70–76.

Jones, E. L., & Forrest, J. D. (1992). Underreporting of abortion in surveys of U.S. women: 1976 to 1988. *Demography, 29,* 113–126.

Katosh, J. P., & Traugott, M. W. (1981). The consequences of validated and self-reported voting measures. *Public Opinion Quarterly, 45,* 519–535.

Keillor, B., Owens, D., & Pettijohn, C. (2001). A cross-cultural/cross-national study of influencing factors and socially desirable response biases. *International Journal of Market Research, 43,* 63–84.

Kiesler, S., & Sproull, L. S. (1986). Response effects in the electronic survey. *Public Opinion Quarterly, 50,* 402–413.

Kitayama, S., Markus, H. R., Matsumoto, H., & Norasakkunkit, V. (1997). Individual and collective processes in the construction of the self: Self-enhancement in the United States and self-criticism in Japan. *Journal of Personality and Social Psychology, 72,* 1245–1267.

Klassen, D., Hornstra, R. K., & Anderson, P. B. (1975). Influence of social desirability on symptom and mood reporting in a community survey. *Journal of Consulting and Clinical Psychology, 43,* 448–452.

Knowles, E. S., & Condon, C. A. (1999). Why people say "yes": A dual-process theory of acquiescence. *Journal of Personality and Social Psychology, 77,* 379–386.

Knowles, E. S., & Nathan, K. T. (1997). Acquiescent responding in self-reports: Cognitive style or social concern? *Journal of Research in Personality, 31,* 293–301.

Krosnick, J. A. (1991). The stability of political preferences: Comparisons of symbolic and non-symbolic attitudes. *American Journal of Political Science, 35,* 547–576.

Krosnick, J. A. (1999). Survey research. *Annual Review of Psychology, 50,* 537–567.

Lai, J., & Linden, W. (1993). The smile of Asia: Acculturation effects on symptom reporting. *Canadian Journal of Behavioural Science, 25,* 303–313.

Lalwani, A. K., & Shavitt, S. (2009). The "me" I claim to be: Cultural self-construal elicits self-presentational goal pursuit. *Journal of Personality and Social Psychology, 97*(1), 88–102.

Lalwani, A. K., Shavitt, S., & Johnson, T. (2006). What is the relation between cultural orientation and socially desirable responding? *Journal of Personality and Social Psychology, 90,* 165–178.

Lamm, B., & Keller, H. (2007). Understanding cultural models of parenting: The role of intracultural variation and response style. *Journal of Cross-Cultural Psychology, 38,* 50–57.

Leary, M.R., & Kowalski, R.M. (1990). Impression management: A literature review and two component model. *Psychological Bulletin, 107,* 34–47.

Lee, C., & Green, R. T. (1991). Cross-cultural examination of the Fishbein behavioral intentions model. *Journal of International Business Studies, 22,* 289–305.

Lee, J. W., Jones, P. S., Mineyama, Y., & Zhang, X. E. (2002) Cultural differences in responses to a Likert scale. *Research in Nursing & Health, 25,* 295–306.

Leite, W. L., & Beretvas, S. N. (2005). Validation of scores on the Marlowe–Crowne Social Desirability Scale and the Balanced Inventory of Desirable Responding. *Educational and Psychological Measurement, 65,* 140–154.

Lenski, G. E., & Leggett, J. C. (1960). Caste, class, and deference in the research interview. *American Journal of Sociology, 65,* 463–467.

Lentz, T. (1938). Acquiescence as a factor in the measurement of personality. *Psychological Bulletin, 35,* 659.

Leung, K., & Bond, M. H. (1980). On the empirical identification of dimensions for cross-cultural comparisons. *Journal of Cross-Cultural Psychology, 20,* 133–151.

Locander, W., Sudman, S., & Bradburn, N. (1976). An investigation of interview method, threat, and response distortion. *Journal of the American Statistical Association, 71,* 269–275.

Locke, J.D., & Baik, K.-D. (2009). Does an acquiescent response style explain why Koreans are less consistent than Americans? *Journal of Cross-Cultural Psychology, 40,* 319–323.

Maheswaran, D., & Shavitt, S. (2000). Issues and new directions in global consumer psychology. *Journal of Consumer Psychology, 9,* 59–66.

Marín, G., Gamba, R. J., & Marín, B. V. (1992). Extreme response style and acquiescence among Hispanics: The role of acculturation and education. *Journal of Cross-Cultural Psychology, 23,* 498–509.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98,* 224–253.

Marlowe, D. and Crowne, D. P. (1961). Social desirability and response to perceived situational demands. *Journal of Consulting Psychology 25,* 109–115.

Marshall, R., & Lee, C. (1998). A cross-cultural, between-gender study of extreme response style. *European Advances in Consumer Research, 3,* 90–95.

Matsumoto, D. (2007). Individual and cultural differences on status differentiation: The status differentiation scale. *Journal of Cross-Cultural Psychology, 38,* 413–431.

McClendon, M. J. (1991). Acquiescence and recency response-order effects in survey interviews. *Sociological Methods & Research, 20,* 60–103.

McQuiston, C., Larson, K., Parrado, E. A., & Flaskerud, J. H. (2002). AIDS knowledge and measurement considerations with unacculturated Latinos. *Western Journal of Nursing Research, 24,* 354–372.

Messick, S. (1991). Psychology and methodology of response styles. In R. E. Snow & D. E. Wiley (Eds.), *Improving inquiry in social science: A volume in honor of Lee J. Cronbach* (pp. 161–200). Hillsdale, NJ: Erlbaum.

Mick, D. G. (1996). Are studies of dark side variables confounded by socially desirable responding? The case of materialism. *Journal of Consumer Research, 23,* 106–119.

Middleton, K. L., & Jones, J. L. (2000). Socially desirable response sets: The impact of country culture. *Psychology & Marketing, 17,* 149–163.

Miller, J. D., Harrel, A. V., & Cisin, I. A. (1986, May). *A new technique for surveying deviant behavior: Item-count estimates of marijuana, cocaine, and heroin.* Paper presented at the annual meeting of the American Association for Public Opinion Research, St. Petersburg, FL.

Minkov, M. (2008). Self-enhancement and self-stability predict school achievement at the national level. *Cross-Cultural Research, 42,* 172–196.

Minkov, M. (2009). Nations with more dialectical selves exhibit lower polarization in life quality judgments and social opinions. *Cross-Cultural Research, 43,* 230–250.

Mirowsky, J., & Ross, C. E. (1991). Eliminating defense and agreement bias from measures of the sense of control: A 2 × 2 index. *Social Psychology Quarterly, 54,* 127–145.

Mitchell, R. E. (1973). Survey materials collected in the developing countries: Sampling, measurement, and interviewing obstacles to intro- and inter-national comparisons. In D. P. Warwick & S. Osherson (Eds.), *Comparative research methods* (pp. 204–226). Englewood Cliffs, NJ: Prentice-Hall.

Moors, G. (2003). Diagnosing response style behavior by means of a latent-class factor approach. Socio-demographic correlates of gender role attitudes and perceptions of ethnic discrimination reexamined. *Quality & Quantity: International Journal of Methodology, 37,* 277–302.

Moscovici, S. (1963). Attitudes and opinions. *Annual Review of Psychology, 14,* 231–260.

Mwamwenda, T. S. (1993). A comparison of two samples, South Africans and Canadians, on social desirability. *Psychological Reports, 72,* 965–966.

Naemi, B. D., Beal, D. J., & Payne, S. C. (2009). Personality predictors of extreme response style. *Journal of Personality, 77,* 261–286.

Narayan, S., & Krosnick, J. A. (1996). Education moderates some response effects. *Public Opinion Quarterly, 60,* 89–105.

Nederhof, A. J. (1985). Methods of coping with social desirability bias: A review. *European Journal of Social Psychology, 15,* 263–280.

Okazaki, S. (2000). Asian American and White American differences on affective distress symptoms: Do symptom reports differ across reporting methods? *Journal of Cross-Cultural Psychology, 31,* 603–625.

Oyserman, D., Coon, H. M., & Kemmelmeier, M. (2002). Rethinking individualism and collectivism: Evaluation of theoretical assumptions and meta-analyses. *Psychological Bulletin, 128,* 3–72.

Paulhus, D. L. (1981). Control of social desirability in personality inventories: Principal-factor deletion. *Journal of Research in Personality, 15,* 383–388.

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology, 46,* 598–609.

Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson & P. R. Shaver (Eds.), *Measures of personality and social psychological attitudes* (pp. 17–59). San Diego, CA: Academic Press.

Paulhus, D. L. (1998a). Interpersonal and intrapsychic adaptiveness of trait self-enhancement: A mixed blessing? *Journal of Personality and Social Psychology, 74,* 1197–1208.

Paulhus, D. L. (1998b). *Paulhus Deception Scales: User's manual.* North Tonawanda, NY: Multi-Health Systems.

Pavlos, A. J. (1972). Racial attitude and stereotype change with bogus pipeline paradigm. *Proceedings of the 80th Annual Convention of the American Psychological Association, 7,* 291–292.

Perera, M., & Eysenck, S. B. G. (1984). A cross-cultural study of personality: Sri Lanka and England. *Journal of Cross-Cultural Psychology, 15,* 353–371.

Phillips, D. L., & Clancy K. J. (1970). Response biases in field studies of mental illness. *American Sociological Review, 35,* 503–515.

Pleck, J. H., Sonenstein, F. L., & Ku, L. (1996). Black–White differences in adolescent males' substance use: Are they explained by underreporting by Blacks? *Journal of Gender, Culture, and Health, 1,* 247–265.

Podsakoff, P. M., MacKenzie, S. M., Lee, J., & Podsakoff, N. P. (2003). Common method variance in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88,* 879–903.

Puntoni, S., & Tavassoli, N.T. (2007). Social context and advertising memory. *Journal of Marketing Research, 44 (2),* 284–96.

Ray, J. J. (1979). Is the acquiescent response style problem not so mythical after all? Some results from a successful balanced F scale. *Journal of Personality Assessment, 4,* 638–643.

Ray, J. J. (1983). Reviving the problem of acquiescent response bias. *Journal of Social Psychology, 121,* 81–96.

Riordan, C., & Vandenberg, R. (1994). A central question in cross-cultural research: Do employees of different cultures interpret work-related measures in an equivalent manner? *Journal of Management, 20,* 643–671.

Ross, C. E., & Mirowsky, J. (1984a). The worst place and the best face. *Social Forces, 62,* 529–536.

Ross, C. E., & Mirowsky, J. (1984b). Socially-desirable response and acquiescence in a cross-cultural survey of mental health. *Journal of Health and Social Behavior, 25,* 189–197.

Rutten, A., De Beuckelaer, A., & Weijters, B. (2008, June). *Between-method convergent validity of three alternative methods to quantify response styles: A case-based comparison.* Paper presented at the International Conference on Survey Methods in Multinational, Multiregional and Multicultural Contexts, Berlin.

Sackeim, H. A., & Gur, R. C. (1979). Self-deception, other-deception, and self-reported psychopathology. *Journal of Consulting and Clinical Psychology, 47,* 213–215.

Schaeffer, N. C. (1980). Evaluating race-of-interviewer effects in a national survey. *Sociological Methods & Research, 8,* 400–419.

Schlenker, B. R., & Britt, T. W. (1999). Beneficial impression management: Strategically controlling information to help friends. *Journal of Personality and Social Psychology, 76,* 559–573.

Schlenker, B. R., Britt, T. W., & Pennington, J. (1996). Impression regulation and management: Highlights of a theory of self-identification. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition,* Vol. 3: *The foundations of social behavior.* New York: Guilford.

Schuman, H., & Converse, J. (1971). The effects of black and white interviewers on black responses in 1968. *Public Opinion Quarterly, 35,* 44–68.

Schuman, H., & Presser, S. (1981). *Questions and answers in attitude surveys: Experiments in question form, wording, and context.* New York: Academic Press.

Schwartz, S. H., Verkasalo, M., Antonovsky, A., & Sagiv, L. (1997). Value priorities and social desirability: Much substance, some style. *British Journal of Social Psychology, 36,* 3–18.

Shapiro, A. H., Rosenblood, L., Berlyne, G. M., & Finberg, J. (1976). The relationship of test familiarity to extreme response styles in Bedouin and Moroccan boys. *Journal of Cross-Cultural Psychology, 7,* 357–364.

Shavitt, S., Johnson, T. P., & Zhang, J. (in press). Horizontal and vertical cultural differences in the content of advertising appeals. *Journal of International Consumer Marketing.*

Shavitt, S., Lalwani, A. K., Zhang, J., & Torelli, C. J. (2006). The horizontal/vertical distinction in cross-cultural consumer research. *Journal of Consumer Psychology, 16,* 325–356.

Si, S. X., & Cullen, J. B. (1998). Response categories and potential cultural bias: Effects of an explicit middle point in cross-cultural surveys. *International Journal of Organizational Analysis, 6,* 218–230.

Sigall, H., & Page, R. (1971). Current stereotypes: A little fading, a little faking. *International Journal of Psychology, 25,* 1–12.

Sigelman, L. (1982). The nonvoting voter in voting research. *Public Opinion Quarterly, 26,* 47–56.

Singelis, T. M., Triandis, H. C., Bhawuk, D., & Gelfand, M. J. (1995). Horizontal and vertical dimensions of individualism and collectivism: A theoretical and measurement refinement. *Cross-Cultural Research: The Journal of Comparative Social Science, 29,* 240–275.

Smith, D. H. (1967). Correcting for social desirability response sets in opinion-attitude survey research. *Public Opinion Quarterly, 31,* 87–94.

Smith, P. B. (2004). Acquiescent response bias as an aspect of cultural communication style. *Journal of Cross-Cultural Psychology, 35,* 50–61.

Smith, P. B., & Bond, M. H. (1998). *Social psychology across cultures* (2nd ed.). Hemel Hempstead, England: Harvester Wheatsheaf.

Smith, P. B., & Fisher, R. (2008). Acquiescence, extreme response bias and culture: A multilevel analysis. In F. J. van de Vijver, D. A. van Hemert, & Y. H. Poortinga (Eds.), *Multilevel analysis of individuals and cultures* (pp. 285–314). New York: Erlbaum.

Smith, T. W. (2003). Developing comparable questions in cross-national surveys. In J. A. Harkness, F. J. van de Vijver, & P. P. Mohler (Eds.), *Cross-cultural survey methods* (pp. 69–91). New York: Wiley.

Soueif, M. I. (1958). Extreme response sets as a measure of intolerance of ambiguity. *British Journal of Psychology, 49,* 329–333.

Steenkamp, J.-B. E. M., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research, 25,* 78–90.

Stening, B. W., & Everett, J. E. (1984). Response styles in a cross-cultural managerial study. *Journal of Social Psychology, 122,* 151–156.

Stöber, J. (3002). The social desirability scale-17 (SDS-17): Convergent validity, discriminant validity, and relationship with age. *European Journal of Psychological Assessment, 17,* 222–232.

Stokes-Brown, A. K. (2006). Racial identity and Latino vote choice. *American Politics Research, 34,* 627–652.

Stricker, L. J. (1963). Acquiescence and social desirability response styles, item characteristics, and conformity. *Psychological Reports, 12,* 319–341.

Sudman, S., & Bradburn, N. M. (1974). *Response effects in surveys.* Chicago: Aldine.

Torelli, C., & Shavitt, S. (in press). Culture and concepts of power. *Journal of Personality and Social Psychology.*

Toner, B. (1987). The impact of agreement bias on the ranking of questionnaire response. *Journal of Social Psychology, 127,* 221–222.

Tourangeau, R., Rips, L., & Rasinski, K. (2000). *The psychology of survey response.* Cambridge: Cambridge University Press.

Tourangeau, R., & Smith, T.W. (1996). Asking sensitive questions: The impact of data collection mode, question format, and question context. *Public Opinion Quarterly, 60,* 275–304.

Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin, 133,* 859–883.

Traugott, M. W., & Katosh, J. P. (1979). Response validity in surveys of voting behavior. *Public Opinion Quarterly, 43,* 359–377.

Traugott, M. W., Traugott, S. & Presser, S. (1992). *Revalidation of self-reported vote* (ANES Technical Report Series, No. nes010160). Ann Arbor, MI: American National Election Studies.

Triandis, H. C. (1972). *The analysis of subjective culture.* New York: Wiley-Interscience.

Triandis, H. C. (1995). *Individualism and collectivism.* Boulder, CO: Westview Press.

Triandis, H. C., Chen, X. P., & Chan, D. K. S. (1998). Scenarios for the measurement of collectivism and individualism. *Journal of Cross-Cultural Psychology, 29,* 275–289.

Triandis, H. C., & Gelfand, M. J. (1998). Converging measurement of horizontal and vertical individualism and collectivism. *Journal of Personality and Social Psychology, 74,* 118–128.

Triandis, H. C., Marín, G., Lisanski, J., & Betancourt, H. (1984). *Simpatia* as a cultural script of Hispanics. *Journal of Personality and Social Psychology, 47,* 1363–1375.

Triandis, H. C., & Triandis, L. M. (1962). A cross-cultural study of social distance. *Psychological Monographs, 76*(21, Whole No. 540).

Tsushima, W. T. (1969). Responses of Irish and Italians of two social classes on the Marlowe-Crowne social desirability scale. *Journal of Social Psychology, 77,* 215–219.

Turk Smith, S., Smith, K. D., & Seymour, K. (1993). Social desirability of personality items as a predictor of endorsement: A cross-cultural analysis. *Journal of Social Psychology, 133*, 43–52.

Van de Vijver, F. J. R., & Leung, K. (1997). *Methods and data analysis for cross-cultural research.* Thousand Oaks, CA: Sage.

Van de Vijver, F. J. R., Ploubidis, G., & Van Hemert, D. A. (2004). *Toward an understanding of cross-cultural differences in acquiescence and extremity scoring.* Paper presented at the Sheth/Sudman Symposium on Cross-Cultural Survey Research Methodology, Urbana, IL.

Van Hemert, D. A., Van de Vijver, F. J. R., Poortinga, Y. H., & Georgas, J. (2002). Structural and functional equivalence of the Eysenck Personality Questionnaire within and between countries. *Personality and Individual Differences, 33*, 1229–1249.

Van Herk, H. (2000). *Equivalence in a cross-national context: Methodological & empirical issues in marketing research.* Unpublished doctoral dissertation, Catholic University, Brabant, The Netherlands.

Van Herk, H., Poortinga, Y. H., & Verhallen, T. M. (2004). Response styles in rating scales: Evidence of method bias in data from six EU countries. *Journal of Cross-Cultural Psychology, 35*, 346–360.

Vernon, S. W., Roberts, R. E., & Lee, E. S. (1982). Response tendencies, ethnicity, and depression scores. *American Journal of Public Health, 116*, 482–495.

Wang, R., Hempton, B., Dugan, J.P., & Komives, S.R. (2008). Cultural differences: Why do Asians avoid extreme responses? *Survey Practice* (published online at: http://surveypractice.org/2008/10/30/cultural-differences/).

Warnecke, R. B., Johnson, T. P., Chávez, N., Sudman, S., O'Rourke, D., Lacey, L., et al. (1997). Improving question wording in surveys of culturally diverse populations. *Annals of Epidemiology, 7*, 334–342.

Warner, S. L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association, 60*, 62–69.

Watkins, D., & Cheung, S. (1995). Culture, gender, and response bias: An analysis of responses to the self-description questionnaire. *Journal of Cross-Cultural Psychology, 26*, 490–504.

Watson, D. (1992). Correcting for acquiescent response bias in the absence of a balanced scale: An application to class consciousness. *Sociological Methods & Research, 21*, 52–88.

Weech-Maldonado, R., Elliott, M. N., Oluwole, A., Schiller, K. C., & Hays, R. D. (2008). Survey response style and differential use of CAHPS rating scales by Hispanics. *Medical Care, 46*, 963–968.

Welkenhuysen-Gybels, J., Billiet, J., & Cambre, B. (2003). Adjustment for acquiescence in the assessment of the construct equivalence of Likert-type score items. *Journal of Cross-Cultural Psychology, 34*, 702–722.

Wells, W. D. (1961). The influence of yea-saying response style. *Journal of Advertising Research, 1*, 5–6.

Welte, J. W., & Russell, M. (1993). Influence of socially desirable responding in a study of stress and substance abuse. *Alcoholism: Clinical and Experimental Research, 17*, 758–761.

Wiggins, J. S. (1959). Interrelationships among MMPI measures of dissimulation under standard and social desirability instructions. *Journal of Consulting Psychology, 23*, 419–427.

Williams, J. E., Satterwhite, R. C., & Saiz, J. L. (1998). *The importance of psychological traits.* New York: Plenum Press.

Wong, N., Rindfleisch, A., & Burroughs, J. E. (2003). Do reverse-worded items confound measures in cross-cultural consumer research? The case of the Material Values Scale. *Journal of Consumer Research, 30*, 72–91.

Yang, Y., Harkness, J. A., Chin T.-Y., & Villar, A. (2010). Response styles in culture. In J. A. Harkness, M. Braun, B. Edwards, T. P. Johnson, L. Lyberg, P. Ph. Mohler, B.-E. Pennell, & T. W. Smith (Eds.), *Survey methods in multinational, multiregional, and multicultural contexts* (pp. 203–223). New York: Wiley.

Yeh, L. L., Kim, K. O., Chompreeda, P., Rimkeeree, H., Yau, N. J. N., & Lundahl, D. S. (1998). Comparison in use of the 9-point hedonic scale between Americans, Chinese, Koreans, and Thai. *Food Quality and Preference, 9*, 413–419.

Ying, Y. (1989). Nonresponse on the Center for Epidemiological Studies-Depression scale in Chinese Americans. *International Journal of Social Psychiatry, 35*, 156–163.

Zax, M., & Takahashi, S. (1967). Cultural influences on response style: Comparisons of Japanese and American college students. *Journal of Social Psychology, 71*, 3–10.