
Original Paper

Survey Response Bias in Comparing Samples from Confucian Asia and North America: Item Type Matters

Robert A. Culpepper, Ph.D.

Professor of Management, Management & Marketing Department,, Nelson Rusche College of Business, Stephen F. Austin State University, PO Box 13070, SFASU, rculpepper@sfasu.edu, (936) 468-1531

Abstract

Research addressing survey response bias among samples from Asian countries with a Confucian legacy, such as China, Korea, and Japan, have been hindered for decades by contradictory theory and inconsistent empirical findings. One theoretical perspective, invoking the Confucian legacy and modesty norms, claims that respondents from these countries choose middle values on Likert-type items, while an opposing perspective, also citing Confucian influences, predicts that respondents from Confucian Asia will choose extreme values. In the latter case rating decisions are said to be formed with reference to a sense of folk knowledge, or what is generally and widely known. Interestingly, each of the diametrically opposed claims have garnered a good deal of empirical support. This paper argues that the disparity found in the literature can be resolved by taking item rating type into account. Previous studies observing the Confucian Asia extreme response pattern have tended to use nomothetic items in which subjects respond to general statements with wide or universal application or about factual information. On the other hand, studies obtaining midpoint or modesty bias largely have employed idiographic items which require a judgement about a particular case, with no general application, and which commonly reference the self, another person such as one's boss, or one's employer. To test this distinction between nomothetic and idiographic ratings, a convenience sample of Chinese and American managers sample was employed. Chinese respondents exhibited significantly higher midpoint scores for items requiring idiographic ratings and significantly higher extreme response scores for nomothetic and factual information-related judgements than did American respondents. This pattern resulted despite item stems being held constant, with respondents asked to make idiographic, nomothetic, and factual/informational ratings referent to the same item stem.

Introduction

Empirical investigations that are international in scope must often bridge large cultural chasms that raise a number of methodological difficulties. Among the most formidable is the pervasive use of Likert-type and semantic differential rating scales and their long-recognized, but under-investigated, susceptibility to culture-related response bias (Adler et al., 1989; Harzing et al., 2012; Jaccard & Wan, 1986; Johnson et al., 2005; Leung & Bond, 1989; Mätus et al., 2012; Zax & Takahashi, 1967). The implications of ignoring response tendency differences can be profound, leading to major inferential errors in cross-cultural research (Cheung & Rensvold, 2000; Chun et al., 1974; Clarke, 2001; Cronbach, 1946; Diamantopoulos et al., 2006; Fischer, 2004; Singh, 1995).

If present, cultural response bias distorts statistical analysis by : (1) rendering group mean differences uninterpretable, (2) spuriously raising or lowering indexes of a measure's internal consistency, (3) spuriously affecting correlations between variables, (4) distorting multivariate causal analyses using correlation-based methods such as regression and canonical correlation analysis, and (5) affecting the results of methods assessing underlying dimensions, such as factor analysis (Arce-Ferrer, 2006; Chun et al., 1974; Diamantopoulos et al., 2006; Peterson et al., 2014).

Researchers recognized long ago that cross-cultural research is perhaps particularly susceptible to error stemming from response bias (Zax & Takahashi, 1967). Unfortunately, the literature addressing cultural differences in response tendencies is still relatively undeveloped, in part because for decades

cross-cultural studies have often ignored this problem, and indeed measurement concerns altogether. Studies addressing response style across cultural samples remain infrequent and have employed student samples almost exclusively (Harzing et al., 2012; van Herk et al., 2004).

Existing research has been largely confined to exploring ways of identifying, classifying, and quantifying response bias, as well as mitigating or at least accounting for it (Cabooter et al., 2017; Cheung & Rensvold, 2000; Dolnicar & Grün, 2007; Fischer, 2004; Mätus et al., 2012; van Herk et al., 2004). The development of theoretical explanations for culture-related response bias has lagged far behind and remains infrequent (Grimm & Church, 1999; Van de Vijver & He, 2014; van Herk et al., 2004). Theoretical progress has been hampered by (1) contradictory theoretical claims and perspectives regarding the cultural basis of response tendencies, (2) conflicting empirical evidence as to the specific nature of response tendencies in particular cultures, and (3) disagreement about whether differences in response tendencies are substantive or mere artifacts of response style.

Interestingly, it is not uncommon in this literature stream to find researchers who assume, as a matter of course, modesty norms for countries have a Confucian legacy, such as China, Korea, and Japan, and more extreme responses for Westerners (e.g., Arce-Ferrer, 2006; Dolnicar & Grün, 2007; Hamamura, Heine, & Paulhus, 2008; Harzing et al., 2012; Mätus et al., 2012), despite this being very much an unsettled question. Such studies routinely cite the work of Chen, Lee and Stevenson (Chen et al., 1995) on this point, and occasionally other studies also observing modesty norms for Confucian Asia (Dolnicar & Grün, 2007; Zax & Takahashi, 1967).

Notwithstanding the evidence that Asian modesty norms play a role in survey response, there is also substantial empirical evidence contradicting this notion. For example, Stening and Everett (Stening & Everett, 1984), widely cited in the literature (e.g., Grimm & Church, 1999; van Herk et al., 2004), demonstrated *extreme* response style (ERS) on the part of respondents from Confucian Asia relative to Australians. In a study comparing several scales across four countries, Yu, Keown and Jacobs (1993) found extreme responses on the part of Chinese, compared to U.S. respondents. Chun et al. (1974) reported that Korean subjects used extreme responses more consistently than did American. In another study comparing Japanese and American management perceptions of managerial strategy principles, researchers (Kotabe et al., 1991) found Japanese unstandardized means higher than American on 24 of 28 items, though they did not analyze extreme and/or midpoint responses. Data were then standardized to avoid bias effects. Adler et al. (1989) encountered such unusual response distributions for Likert-type items among Chinese managers that they abandoned the intended substantive inquiry addressing Chinese-American differences in managerial attitudes -- analyzing instead methodological barriers to cross-cultural research. Finally, in the decision-making literature, Yates and colleagues (Yates et al., 1997), citing the Stening and Everett study mentioned above, explored whether the phenomenon of “Asian Overconfidence” might be related to extreme response style on the part of Asian respondents.

This remarkable disparity in empirical evidence on the response bias question is not merely that some studies show responses from Confucian Asia to be more extreme than for Westerners, while the rest show no meaningful differences. Rather, findings directly contradict each other - some studies show Asian responses more extreme and others show Western responses more extreme. The resolution of this conundrum doubtless depends on further theoretical exploration and development. In the meantime, extreme response style (ERS) and midpoint response bias (MRS) remain perhaps the least understood among response biases (Peterson et al., 2014).

Theoretical explanations for ERS and MRS, thus far, are also varied and no less contradictory than the empirical evidence. A number of studies have offered Individualism (Hofstede, 1984) as an explanation, suggesting that more individualistic Westerners express stronger and more extreme opinions than relatively collectivist Asians, who emphasize interpersonal harmony (Chen et al., 1995; Johnson et al., 2005). Researchers have also invoked Individualism/Collectivism to explain response bias differences across countries—in one instance, European countries (Johnson et al., 2005). Unfortunately, no clear pattern of empirical evidence for the influence of Individualism on response bias has emerged (Grimm & Church, 1999; Van Dijk et al., 2009; van Herk et al., 2004).

Confucian-based modesty norms (Chen et al., 1995; Hui & Triandis, 1989) have also been used to

explain higher midpoint response and lower extreme response for East Asian cultures with a Confucian legacy--such as China, Taiwan, Singapore, Japan, and Korea – as compared to Western respondents. In such cultures, the argument goes, assertiveness and displaying of strong, independent opinions do not constitute virtues, but rather are frowned upon. As discussed above, research findings have not always squared with Asian modesty in survey-taking--and in some cases have supported more extreme Asian responses (Adler et al., 1989; Kotabe et al., 1991; Stening & Everett, 1984; Yu et al., 1993).

Remarkably, some researchers ignore the arguments by Chen et al. (1995) and others about the Confucian legacy as it relates to survey response, and suggest that this history should lead us to expect more *extreme* responses with respect to Confucian Asia (Yates et al., 1997). In this latter view, Confucian cultures are less adversarial and debate between opposing viewpoints is considered undesirable (Yates & Lee, 1996). The preferred means of acquiring knowledge is not through debate over competing ideas, but through attaining an understanding of what is already known and accepted as correct. Learning is seen primarily as emulation, the acquisition of ideals, practices and procedures that have general acceptance and have stood the empirical test of time. Thus, decision-makers from Confucian cultures are said to revert to a sense of that which seems time-tested and widely accepted, rather than constructing and weighing arguments on both sides of an issue. Indeed, some studies have supported the recruitment of fewer arguments by Chinese subjects in decision-making (Lee et al., 1995; Li et al., 2009, 2011). Similarly, survey respondents in such cultures should be less prone to weigh relevant pros and cons in their responses--referencing instead a sense about a proposition's consistency with relevant accepted thinking or traditional folk knowledge (Yates et al., 1997). In so doing, they are less prone to choose middle values on semantic differential or Likert-type items that would imply ambivalence or compromise between conflicting rationales or bases of support, and more likely to choose extreme values instead.

Indeed, a number of studies in the decision-making literature have found that respondents from Confucian Asia, such as the Chinese, to be more overconfident in their judgments and evaluations, relative to Westerners (Whitcomb et al., 1995; Wright & Phillips, 1980; Yates et al., 1989, 1990, 1990). These studies typically have compared cultures based on the level of confidence subjects have in their judgements relative to the accuracy of those judgements. Worth noting is that this literature attributes overconfidence in judgements to cultural/cognitive differences among respondents, i.e. to substantive differences (Yates et al., 1997), rather than to a mere artifact of decisional or survey-taking "style" (Chun et al., 1974)--which is far less theoretically accessible.

To summarize, the literature regarding culture-based differences in ERS and MRS contains contradictory theoretical claims and conflicting empirical evidence. Currently, opposing theoretical claims about MRS AND ERS share a common feature--they incorporate "one size fits all" explanatory reasoning. In each case, the cultural feature purportedly leading to response bias--be it Individualism, Asian modesty norms, or the Confucian non-deliberative/folk knowledge orientation--is presented, at least implicitly, as operating independently of differences in such factors as rating task and item content. The current study investigates a potential mitigating factor that may shed light on the longstanding theoretical impasse--namely, the role that survey item type plays, and specifically, the nature of the rating judgement that is required for a given item.

The Role of Item Content and Rating Type

Previous studies, even when focused entirely on the question of response bias, have largely ignored the nature of rating judgements--and indeed item characteristics altogether. A few studies have examined the role of response scale length, e.g., three-point versus five-point, with some work claiming higher observed ERS when using longer scales (Hui & Triandis, 1989; Jin & Wang, 2014) and some work supporting a *reduction* in ERS for longer scales, reaching up to seven points (Clarke, 2001). A few researchers have addressed the role of item domain (Arce-Ferrer, 2006; Cabooter et al., 2017; Van Dijk et al., 2009). While theory relating item domain to response tendencies generally has been absent, Van Dijk (2009) observed that "personally relevant" items generated greater differences in response bias across European countries and Arce-Ferrer (2006) suggested that item content related to face-saving, may have explained rural-urban differences in ERS.

Despite these exceptions, the assumption that item rating task and content make no difference has been

so widespread that a common and recommended practice for identifying response bias in a given study has been to extract individual items from multiple scales to form a composite ERS measure. The point of this practice is to measure ERS without risking conflation with substantive score variance that might otherwise result from using same-scale items as an index. With this purpose in mind, for example, Greenleaf developed an ERS measure (1992) for use in other studies, by deliberately incorporating items from completely different content domains.

Despite the past practice of using divergent items to assess ERS, there is growing evidence that that cultural traits interact with item content in affecting response bias--and that item content matters for the level and type of response bias observed (Arce-Ferrer, 2006; Cabooter et al., 2017; De Jong et al., 2008; Ross & Mirowsky, 1984; Van Dijk et al., 2009). For example, it seems reasonable to expect that modesty norms, whether arising from Confucian history or Collectivist norms, operate primarily *when items are personal in nature*, i.e., when they relate to one's sense of personal identity. That is, modesty should be triggered when items ask about one's tastes, preference, opinions, performance--or when items address groups related to personal identity, such as family, company, or other affiliations. But it is not at all clear that modesty norms should extend beyond identity-related cases--as in the case of items assessing widely applicable propositions or principles, or items soliciting factual information with low identity relevance. It is interesting to note that one study from the 1980s opined, as if stating the obvious, that Japanese modesty norms do not apply to expressing opinions about factual matters (Horler & Yamazaki, 1986).

While modesty norms may fail to influence scores for items addressing general propositions or factual content relatively devoid of identity-related content, the folk knowledge rationale advanced by Yates and colleagues (Yates et al., 1997; Yates & Lee, 1996) seems much more germane. Specifically, respondents in Collectivist cultures with strong conformity norms seem unlikely to feel inhibited about expressing agreement with propositions or principles seen as popular in appeal or in line with normative folk knowledge. Respondents in such cultures should be inclined, if anything, toward the extreme responses posited by Yates et al. Though Yates and Lee make their argument based on the Confucian cultural legacy, limited to certain East Asian cultures, it may well apply beyond, due to conformity norms found more generally in Collectivist cultures (Arce-Ferrer, 2006; Hofstede, 1984).

To summarize, though Yates and colleagues' folk wisdom argument may explain extreme response bias for respondents in Confucian Asia in reacting to generalized propositions with relatively universal application (hence "nomothetic"), this argument makes little sense when applied to items characterizing the self or specific groups, such as one's family or company--because there are typically no societal referents in these cases. While the folk wisdom argument may well apply to rather to generalities, principles, and truisms, it cannot apply to items requiring a judgment about a specific case (hence "idiographic"), whether self-referencing or not.

This nomothetic versus idiographic distinction has been widely used across different disciplines and contexts within the social sciences, dating back to an early usage by the psychologist Gordon Allport (1937). It would appear to have significant utility here in clarifying theory relevant to cross-cultural differences in response bias. Specifically, existing empirical evidence suggests that observed response tendencies vary substantially across nomothetic and idiographic item categories. All of the above-mentioned cross-cultural studies reporting ERS for samples from Confucian Asia (Adler et al., 1989; Chun et al., 1974; Kotabe et al., 1991; Stening & Everett, 1984; Yu et al., 1993) used nomothetic items. Obversely, studies finding higher extreme response for Western samples relative to those from Confucian Asia--or greater midpoint response for samples from Confucian Asia than Western samples--have employed idiographic items (Chen et al., 1995; Dolnicar & Grün, 2007; Wang et al., 2008; Zax & Takahashi, 1967). The literature on cross-cultural differences also contains studies suggesting, however tacitly, that response tendencies occur in line with nomothetic-idiographic differences in item or rating scale content, without such studies reporting or even raising the question of response bias. For example, one study comparing American and Taiwanese school children found no differences for three attitudinal scales containing idiographic items, but found the Taiwanese mean higher for the scale composed exclusively of nomothetic items (Stigler et al., 1985)--in the latter case indicative of possible ERS. In these cases, one suspects that item type has played a role, but without conducting a new analysis with the original data, it is often impossible to tell.

The use of nomothetic versus idiographic items may even apply to differences in response style observed intra-culturally. For example, Greenleaf's ERS measure (1992) interspersed both idiographic items (e.g., "I try to avoid foods that are high in cholesterol", "I am a homebody" "I work very hard most of the time") and nomothetic items (e.g., "Everyone should use mouthwash to control bad breath", "A college education is very important for success in today's world", "Investing in certificate deposits is too risky for most families"). Despite the Greenleaf scale's rationale that mixed item content is superior for measuring ERS, the above-mentioned Arce-Ferrer study (2006) demonstrated that item-level extreme response indices were, on average, 47 percent higher for nomothetic items (.56) than for idiographic items (.38) in Mexican samples spanning urban and rural settings. Moreover, ERS scores were uniformly higher for rural respondents than urban, across all items; Arce-Ferrer cited stronger collectivist norms in rural areas as the reason.

Additionally, studies comparing European countries suggest that the idiographic-nomothetic distinction may have relevance beyond Asian-Western comparisons. For example, Diamantopoulos et al. (2006) showed that the use of first person (idiographic) versus third person (nomothetic) item stems impacted midpoint and extreme responding levels within countries and between countries. In addition, Cabooter et al. (2017), found a domain-specific ERS component in addition to a general ERS component in factor analytic results. The political item domain had higher loadings on the general ERS factor than items in the domains of consumer behavior and interpersonal relationships. However, the high-ERS political domain was assessed exclusively with nomothetic items (e.g., "Students should be trained in times of peace to carry out military duties") while the consumer behavior domain (e.g., "Buying a high-price brand makes me feel good about myself") and interpersonal relationship domain (e.g., "I am very happy with my friendships") were tapped using overwhelmingly—more than 80 %--idiographic and personally relevant items. In light of the above discussion, findings by Cabooter et al. may have stemmed from nomothetic-idiographic differences in items type rather than from the topic-related domain differences reported in the study.

Despite significant attention to cross-cultural differences in ERS and MRS, relatively few studies have sought to explain them, and most of these have used student samples (van Herk et al., 2004). Accordingly, conceptual progress in this area has been modest (Van de Vijver & He, 2014). The current study, rather than provide a confirmation of theoretical claims in the literature, seeks rather to examine a potential moderating factor that determines which response biases result in comparisons between survey respondents from Confucian Asia compared to those from North America. This investigation is informed by theory, but does not seek to support one theoretical position over another – but rather to provide practical guidance in the types of response biases to expect and under which conditions. This study identifies one area of contradictory findings and seeks to provide empirical evidence to help in resolving this contradiction. The current line of inquiry focused on whether the type of item rating--idiographic, nomothetic, or factual/informational--yielded different patterns of response bias. The use of nomothetic item ratings or factual item ratings were expected to yield higher ERS and lower MRS scores for respondents from Confucian Asia than for Western, while idiographic items were expected to result in higher ERS scores and lower ERS scores for Western respondents than those from Confucian Asia.

The current study employed a convenience sample of Chinese and American managers, which is preferable to student samples, commonly used in testing theoretical explanations (Arce-Ferrer, 2006; Harzing et al., 2012), Student samples are not adequately representative of larger populations, usually lack life experience substantial life or work experience, and fail to move us any closer to discerning boundary conditions that may exist for the findings in question.

The nature of this study's convenience sample and survey helped to avoid the potentially confounding variable of whether the topical domain of items was determinative in response tendency. Items within a given domain can be formulated in either an idiographic or nomothetic manner. For example, "it is important for a person to have self-esteem" represents a nomothetic format, while the item "I have a high level of self-esteem" exemplifies an idiographic format. Nevertheless, both items reference self-esteem and are from the same content domain. Typically, item format derives from the wording of an item's stem or the labeling of anchors on the corresponding rating scale.

The current sample afforded the advantage of a rather novel item type--respondents were required to perform three different rating tasks related to the same item stem. Each respective item stem was accompanied by three Likert-type scales that required, respectively, (1) a rating/judgement that was factual/informational in nature, (2) a nomothetic rating indicating one's assent to a proposition with general application, and (3) a rating requiring an evaluation of a specific case. Item examples are given in the Appendix. The unusual item format in this study afforded a unique opportunity to not only hold the general item domain constant across empirical tests, but to hold the specific substantive focus of each respective item constant, while simultaneously varying the nature of rating tasks--factual, nomothetic, and idiographic. The primary purpose of the current study was to determine whether response tendencies, namely ERS and MRS, varied across the three different types of rating tasks.

Method

Sample

To avoid the use of problematic student samples (e.g., Harzing et al., 2012; (van Herk et al., 2004)), the current study employed a convenience sample of managers employed in American-Chinese joint ventures. A total of 198 useable surveys were obtained, 109 of whom were from Chinese respondents situated in the PRC, and 89 of whom were Americans based in the US.

Instrument

Surveys employed 33 item stems listing various business performance criteria across several areas--financial (e.g., "return on investment"), marketing (e.g., "customer brand awareness"), operational (e.g., "manufacturing quality control"), technological (e.g., "product design capability") and relational (e.g., "relations between the joint venture's parent firms"). For each item stem describing a specific performance dimension, subjects were asked to rate on a seven-point scale, (1) its frequency of usage in assessing performance (thus representing a factual/informational rating task asking about specific knowledge), (2) its general importance as a performance dimension (thereby indicating assent to a nomothetic proposition, i.e., the importance of the dimension), and (3) the perceived actual performance of the joint venture on that particular dimension (representing an idiographic rating task involving a specific case, i.e., the particular joint venture in which the respondent was employed). Semantic anchors for the usage frequency of each performance dimension were "Never" = 1, "Regular" = 4, and "Frequent" = 7. For each performance dimension's importance, anchors were "Minimal" = 1, "Moderate" = 4, and "Great" = 7. Anchors for the joint venture's actual performance on respective dimensions were "Poor" = 1, "Okay" = 4, and "Excellent" = 7. All items were translated from English to Chinese and then independently back-translated according to the method recommended by Brislin (1970).

Data Analysis

Extreme response score (ERS) and midpoint (MRS) response scores were calculated according to the Bachman and O'Malley protocol (Bachman & O'Malley, 1984), the most widely used response bias indexing method (Peterson et al., 2014). An "extreme response" score, ERS, was calculated for each respondent indicating the number of times a "1" or "7" response was chosen and a "midpoint response" score, MRS, computed for the number of "4"s selected. In addition to ERS and MRS, midrange response scores were computed for respondents who selected "3", "4", or "5." Multivariate analysis of variance (MANOVA) was conducted to control Type I error, which is particularly problematic when analyzing a series of between-group mean differences in response frequencies. The chance of observing a significant difference between groups for multiple dependent variables using difference tests for each outcome, assuming no true population differences, increases according to the number of tests performed. MANOVA controls type I error by testing all dependent variables simultaneously for significant differences across independent variable categories. This procedure included six dependent variables, namely, respondents' ERS and MRS scores for each of the three scales measuring usage frequency, importance, and actual performance. Respondents' nationality served as the independent variable.

In the current study, the focus went beyond whether response bias varied by culture and addressed whether such bias differed within culture and across rating tasks. Specifically, nomothetic and

informational/factual rating tasks, were expected--based on the Confucian non-deliberative/folk knowledge orientation--to elicit more extreme responses and fewer midpoint responses by Chinese respondents. The idiographic rating task was expected, based on Confucian modesty norms, to produce the opposite response pattern, with fewer extreme responses and more midpoint responses for Chinese respondents. As discussed above, the investigation of response bias contingent on rating task appears to be a promising line of theory development to help clarify conflicting findings in the research literature related to culture-related response bias. To explore this possibility, univariate F tests were computed in connection with MANOVA to determine whether rating task played a role in any observed country differences in response tendencies.

Results

MANOVAs indicated overall differences ($p < .01$, $p < .001$) between Chinese and U.S. respondents in their preferences for extreme and midpoint response scores (see Table 1). Consistent with the non-deliberative/Confucian folk knowledge rationale discussed above, extreme scores were selected significantly more frequently by Chinese respondents than American for nomothetic evaluation tasks ($p < .01$; see Figure 1) and factual knowledge rating tasks ($p < .001$), but significantly less frequently for idiographic evaluation tasks ($p < .001$).

Also consistent with folk knowledge theory, midpoint scores were chosen less frequently by Chinese respondents ($p < .001$; see Figure 1) for factual knowledge and nomothetic evaluation tasks ($p < .001$). Contrary to expectations, however, Chinese managers did not assign significantly more midpoint scores than American for idiographic evaluation items. Nonetheless, the analysis of *midrange* scores obtained the expected pattern, with Chinese respondents choosing fewer midrange values for factual ($p < .001$) and nomothetic ($p < .001$) items, and significantly more midrange values for the idiographic task ($p < .01$).

Discussion

The results of the current study provide evidence relating to several questions regarding survey response bias. First, current findings add to existing evidence that the tendency toward extreme responses is substantive, culturally-influenced, and not a mere artifact of a test-taking style (Chun et al., 1974; Yates et al., 1997). A test-taking style would apply uniformly across item types. Second, current findings--in line with several previous studies (Adler et al., 1989; Kotabe et al., 1991; Stening & Everett, 1984; Yu et al., 1993)--challenge assumptions by some researchers that a generalized modesty effect for survey responses in Confucian Asia is well-established. The modesty supposition is understandable given its intuitive appeal and the fact that it squares well with Hofstede's influential work (1984, 2001) regarding the Individualism-Collectivism dimension, in particular. However, current evidence suggests that modesty norms do not apply across survey item types. In the current sample, respondents from Confucian Asia exhibited *higher ERS* than Western respondents under two of the item rating conditions.

Third, the most important implication of the current study is that the apparent contradiction in empirical findings in the response bias literature may be resolved by taking rating type into account. In other words, *both* sides in the seemingly contradictory claims about cross-cultural differences in ERS and MRS biases may be correct. Specifically, current results suggest that samples from Confucian Asia may well exhibit both more extreme responses and more middling responses than Western samples--even within the same survey--depending on the nature of survey items. Chinese ERS scores were higher and MRS scores lower than American when respondents were faced with the more abstract and principle-oriented, nomothetic rating type and the more fact-oriented, informational rating task--consistent with theory by Yates and colleagues (Yates et al., 1997; Yates & Lee, 1996). Remarkably, the pattern completely reversed for idiographic items, with the Chinese sample showing *lower ERS* and higher midrange scores--in line with theory relating to modesty bias in Collectivist cultures, and Confucian cultures in particular.

As discussed above, the current study makes explicit a pattern that already exists, albeit implicitly, in the literature. Numerous previous studies tacitly support the current findings that nomothetic rating tasks yield higher extreme response tendency and/or lower midpoint response tendency for samples

from Confucian Asia (Adler et al., 1989; Chun et al., 1974; Kotabe et al., 1991; Stening & Everett, 1984; Yu et al., 1993). A number of studies suggest that the same pattern occurs when a survey requires judgements about factual information (Whitcomb et al., 1995; Wright & Phillips, 1980; Yates et al., 1990). Also reflected in past work is the reversal of this pattern for idiographic ratings that often reference the self or one's organization, leading to lower extreme response and/or higher midpoint response for respondents from Confucian Asia relative to Western samples (Chen et al., 1995; Dolnicar & Grün, 2007; Wang et al., 2008; Zax & Takahashi, 1967). This body of previous work provides a measure of assurance that current results are not one-off and data-specific.

The divergence in theoretical claims and findings in the literature examining the role of culture in response--and the relative dearth of attention given to this divergence—is surprising. This state of affairs serves to illustrate the ease with which culture-based arguments can be marshalled in support of inconsistent and even opposing claims. For example, in the decision-making literature, Yates and colleagues have employed a culture-based argument (Yates et al., 1989; Yates & Lee, 1996) to explain why subjects from Confucian Asia use more extreme responses than Western subjects, while Hamamura et al. (2008) used a cultural rationale in support of why these particular Asian countries should use *fewer* extreme responses. Even more striking is that *similar* arguments at times are marshalled in the service of *opposing* claims. Yates and colleagues hold that Westerners, unlike those from Confucian Asia, weigh opposing rationales and compromise toward a scale's middle values to resolve the contradiction. Hamamura et al. argue rather that those from Confucian Asia, based on a trait called dialectical thinking, look at opposing rationales and honor the truth in each, thus inclining them toward middle values--while Westerners cannot abide the contradiction and resolve toward one end of the scale or the other. The situation becomes even more conflicted given that each group of researchers has obtained evidence for their respective and mutually exclusive claims.

Results from the current study suggest that the Yates-Hamamura contradiction may be resolved by taking rating type into account. The Asian Overconfidence literature (Li et al., 2006; Yates et al., 1996) argues for extreme response tendency on the part of respondents from Confucian Asia, but also happens to use primarily fact-related responses as the criterion, thus leading to findings of extreme responses for those from Confucian Asia and middling responses for Westerners. Hamamura et al. (2008) used idiographic, self-referencing items from the Rosenberg (1965) self-esteem scale and came to the opposite conclusion--finding Western responses more extreme and Asian responses more middling. Thus, by taking item rating type into account--it may be possible to reconcile the opposing theoretical rationale's by specifying boundary conditions for each explanation. Specifically, theory proposed by Yates, et al. should apply in cases of fact-based survey items, and in all likelihood, nomothetic items as well, while that by Hamamura et al. should apply in studies employing idiographic items. An important caveat here is that, insofar as the current convenience sample did not allow for measuring culture at the individual level, the current results fall short of providing robust tests of theoretical claims. This limitation notwithstanding, current results suggest that item type makes a difference in response bias patterns, important in and of itself, regardless of any related theoretical reasoning, and provides guidance about possible fruitful directions for tests of theoretical claims.

Many cross-cultural studies mix rating types, sometimes unavoidably so, by selecting certain pre-existing scales for inclusion. Nevertheless, in light of current findings, the implications of mixing item types in a survey may be problematic from a response bias perspective. In particular, the practice of creating ERS indices from mixed rating types--whether done on an ad hoc basis using items drawn from that study's employed scales, or by employing a ready-made ERS index such as the Greenleaf scale (1992)--appears likely to obfuscate the true direction and/or magnitude of response bias.

Current findings further suggest that in addition to the problems stemming from ERS indices mixing topical domains (Cabooter et al., 2017), scores from a given ERS index may be dramatically skewed based on proportion of rating types included. The Greenleaf (1992) ERS scale, for example, mixes identity-related idiographic items with nomothetic items. The Greenleaf index should thus understate or overstate a given scale's true ERS according to the proportion of item types in the study's included scales and which culture the study's respondents represent.

For example, in cross-cultural studies comprised primarily or exclusively of idiographic items, quite

common in the psychological literature, the Greenleaf scale (1992) should impute exaggerated ERS to survey responses in a sample from Confucian Asia. This is due to the 16-item index's inclusion of five nomothetic items, inflating ERS scores in these cultures. In other words, a bias index composed entirely of items from the same rating type, i.e., in this case idiographic items, would yield a truer picture of culture-related bias specific to this item type--in this case, lower ERS and perhaps higher MRS. The Greenleaf scale includes enough nomothetic items to generate a substantially higher and spurious ERS score in East Asian samples. Other concerns for the Greenleaf scale are that it was developed using an exclusively North American sample and mixes items' topical domains--thus, its utility across cultures is questionable in any event. In addition to ready-made scales such as the Greenleaf scale, ad hoc bias indices pulled from across scales within the same study should be susceptible to the same skewing, the direction of which would depend on which rating types were represented in the index--and in what proportions.

This study was exploratory in nature and subject to certain limitations. First, though existing theory helped in providing guidance for how item type might factor into response bias in comparisons between North American and Confucian Asia, the lack of individual-level measures of cultural attributes prevented robust support of theoretical propositions. The convenience sample provided a more realistic setting and focus for tests, but precluded the tailoring of the survey to questions of interest. Second, the current study used an unusual item format employed in the convenience sample. Although this format allowed for holding item content completely constant while varying rating type, it is highly atypical in cross-cultural studies. This raises the question of whether results are representative of what might be found using other, more typical Likert-type items, or other variable-response formats. However, this uncommon item format likely represents a *conservative* test of the role of rating type, insofar as the matching of a common item stem to multiple response sets should mitigate toward *similarity* across the three response sets, rather than the obtained disparity. Advantages of the employed item format notwithstanding, this format does not appear essential to any future work investigating the effect of rating judgement type. As discussed above, differences in response bias across nomothetic, idiographic, and factual categories have been quite evident in previous studies employing the more typical variable-response items.

Third, the idiographic rating tasks in this convenience sample involved more than a particularized assessment of a specific case, but were conflated with aspects of personal identity--namely, these items asked about the performance of one's own company. The use of identity-related, idiographic items is common in fields such as psychology and marketing, and current results may be representative for studies in which such items are employed. The current study did not bear specifically on the question of whether East Asian responses to such idiographic items were due solely to the idiographic, single-case, nature of items or to the self-referencing aspect of items. In either case, modesty norms invoked in the literature could be in play, since cultural modesty may apply more broadly, without being confined to self-referencing items. In other words, cultural modesty may mitigate toward more caution in making single-case assessments where one is out on a limb, so to speak, and potentially out of step with one's sense of collective, folk sensibilities. Nonetheless, further work is needed to investigate whether the same response bias results obtain when idiographic items are low in identity relevance.

Though the interaction of item rating type and response bias demonstrated in this study corresponds to patterns evident in previous work, more work is needed to determine whether this pattern can be detected prospectively in comparisons involving other East Asian samples, and more generally, comparisons addressing other regions of the world considered culturally collectivist or high context. In addition to cross-cultural comparisons, future work might examine whether the nomothetic-idiographic item contingency accounts for *intra*-cultural differences in ERS and MRS levels, as response patterns in the Arce-Ferrer (Arce-Ferrer, 2006) study implicitly suggest in obtained score differences between urban and rural respondents. Future research might also explore other boundary conditions such as whether the current sample of managers accounted for this study's results. Interestingly, the current finding that East Asian managers exhibit ERS when confronted with nomothetic items mirrors findings from earlier work also using samples of East Asian managers (Adler et al., 1989; Kotabe et al., 1991). These suggest that Western managers are more reticent to express strong agreement with more abstract, generalized propositions, while East Asian managers have no such resistance, although samples

comprised by those employed in other lines of work might not yield the same pattern.

Current findings also raise the question of whether the lines between idiographic, nomothetic, and factual items will always be clear and whether survey items, in some cases, may defy distinct categorization. Nevertheless, the current item categories appear to be useful in differentiating survey items employed across a variety of previous studies in which response bias was a concern. Future work may provide further delineation, elaboration or modification of these categories.

It is worth noting that, due to many studies' sparse or completely absent item descriptions, it is often impossible to determine whether the items used were idiographic, nomothetic, or factual. In addition, studies investigating response bias sometimes fail to make it clear whether items are self-referencing or not. A common practice is to describe the topical domain and give a couple of sample items--thus preventing an adequate assessment of the role that item format, content, and rating type may have played. Most problematic of all, some studies fail to give any account, whatsoever, of the types of items used. For example, one major study reporting ERS levels across 26 countries merely indicated that all items came from two global marketing research companies (De Jong et al., 2008). Given the fact that items were marketing-related, typically idiographic and self-referencing, one might reasonably surmise--based on the current review of past work and current results--that the very low levels of ERS obtained for the Chinese and Taiwanese obtained by De Jong et al. were due largely to idiographic item content and rating tasks.

Finally, the current investigation went beyond questions of response format length and item domain, and may prove helpful in reconciling findings which have diverged for decades. The type of judgement or rating appears to matter, a factor that is largely independent of item content. Current results comport with those of Diamantopoulos et al. (2006) suggesting that cross-national differences in response bias are not stable across changes in stimulus format, such as item wording, scale type, and response categories. However, the implications of this study go further – and suggest that cross-country differences in response bias potentially interact with item format in a systematic and predictable manner. The current study offers a theoretical explanation that helps make sense of response bias findings obtained within and across cross-cultural studies. It also provides a clarification of which differences in response style to expect across cultures and under what conditions--and helps sort out the puzzling array of empirical and theoretical inconsistencies endemic to this literature.

Author Note

I would like to thank Liming Zhao, Lexmark International, Inc. for his help in providing the study's data and for useful suggestions on an earlier version of this paper. Thanks also go to Clyde Lowery, Lockheed Martin Corporation, for his help supporting earlier drafts of this paper.

References

- Adler, N. J., Campbell, N., & Laurent, A. (1989). In Search of Appropriate Methodology: From Outside The People's Republic of China Looking In. *Journal of International Business Studies*, 20(1), 61–74. <https://doi.org/10.1057/palgrave.jibs.8490351>
- Allport, G. W. (1937). *Personality: A psychological interpretation*. New York: Henry Holt and Co.
- 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(3), 374–392. <https://doi.org/10.1177/0013164405278575>
- Bachman, J. G., & O'Malley, P. M. (1984). Yea-saying, nay-saying, and going to extremes: Black-white differences in response styles. *Public Opinion Quarterly*, 48(2), 491–509.
- Brislin, R. W. (1970). *Back-Translation for Cross-Cultural Research—Richard W. Brislin, 1970*. <https://journals.sagepub.com/doi/abs/10.1177/135910457000100301>
- Cabooter, E., Weijters, B., De Beuckelaer, A., & Davidov, E. (2017). Is extreme response style domain specific? Findings from two studies in four countries. *Quality & Quantity*, 51(6), 2605–2622. <https://doi.org/10.1007/s11135-016-0411-5>

- Chen, C., Lee, S., & Stevenson, H. W. (1995). Response Style and Cross-Cultural Comparisons of Rating Scales Among East Asian and North American Students. *Psychological Science*, 6(3), 170–175. <https://doi.org/10.1111/j.1467-9280.1995.tb00327.x>
- 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(2), 187–212. <https://doi.org/10.1177/0022022100031002003>
- 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(4), 465–480. <https://doi.org/10.1177/002202217400500407>
- Clarke, I. (2001). Extreme response style in cross-cultural research. *International Marketing Review*, 18(3), 301–324.
- Cronbach, L. J. (1946). Response Sets and Test Validity. *Educational and Psychological Measurement*, 6(4), 475–494. <https://doi.org/10.1177/001316444600600405>
- De Jong, M. G., Steenkamp, J.-B. E., 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(1), 104–115.
- 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(8), 925–935.
- Dolnicar, S., & Grün, B. (2007). Cross-cultural differences in survey response patterns. *International Marketing Review*, 24(2), 127–143. <https://doi.org/10.1108/02651330710741785>
- 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(3), 263–282. <https://doi.org/10.1177/0022022104264122>
- Greenleaf, E. A. (1992). Measuring extreme response style. *Public Opinion Quarterly*, 56(3), 328–351.
- Grimm, S. D., & Church, A. T. (1999). A Cross-Cultural Study of Response Biases in Personality Measures. *Journal of Research in Personality*, 33(4), 415–441. <https://doi.org/10.1006/jrpe.1999.2256>
- Hamamura, T., Heine, S. J., & Paulhus, D. L. (2008). Cultural differences in response styles: The role of dialectical thinking. *Personality and Individual Differences*, 44(4), 932–942.
- Harzing, A.-W., Brown, M., Köster, K., & Zhao, S. (2012). Response Style Differences in Cross-National Research. *Management International Review*, 52(3), 341–363. <https://doi.org/10.1007/s11575-011-0111-2>
- Hofstede, G. (1984). *Culture's Consequences: International Differences in Work-Related Values*. SAGE.
- Hofstede, G. (2001). *Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations*. SAGE Publications.
- Horler, F., & Yamazaki, J. (1986). Response styles of Japanese and American college students. *JALT J*, 8, 83–90.
- Hui, C. H., & Triandis, H. C. (1989). Effects of Culture and Response Format on Extreme Response Style. *Journal of Cross-Cultural Psychology*, 20(3), 296–309. <https://doi.org/10.1177/0022022189203004>
- Jaccard, J., & Wan, C. K. (1986). Cross-Cultural Methods for the Study of Behavioral Decision Making. *Journal of Cross-Cultural Psychology*, 17(2), 123–149. <https://doi.org/10.1177/0022002186017002001>
- Jin, K.-Y., & Wang, W.-C. (2014). Generalized IRT Models for Extreme Response Style. *Educational and Psychological Measurement*, 74(1), 116–138. <https://doi.org/10.1177/0013164413498876>

- Johnson, T., 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(2), 264–277.
- Kotabe, M., Duhan, D. F., Smith, D. K., & Wilson, R. D. (1991). The Perceived Veracity of PIMS Strategy Principles in Japan: An Empirical Inquiry. *Journal of Marketing*, 55(1), 26–41. <https://doi.org/10.1177/002224299105500103>
- Lee, J.-W., Yates, F. J., Sninotsuka, H., Singh, R., Onglatcc, M. L. U., Yen, N., Gupta, M., & Bhatnagar, D. (1995). *Cross national difference in overconfidence*. <http://vsilir.iima.ac.in:8080/xmlui/handle/11718/9381>
- Leung, K., & Bond, M. H. (1989). On the Empirical Identification of Dimensions for Cross-Cultural Comparisons. *Journal of Cross-Cultural Psychology*, 20(2), 133–151. <https://doi.org/10.1177/0022022189202002>
- Li, S., Bi, Y.-L., & Rao, L.-L. (2011). Every Science/Nature Potter Praises His Own Pot—Can We Believe What He Says Based on His Mother Tongue? *Journal of Cross-Cultural Psychology*, 42(1), 125–130.
- Li, S., Bi, Y.-L., & Zhang, Y. (2009). Asian Risk Seeking and Overconfidence 1. *Journal of Applied Social Psychology*, 39(11), 2706–2736.
- Li, S., Chen, W.-W., & Yu, Y. (2006). The Reason for Asian Overconfidence. *The Journal of Psychology*, 140(6), 615–618. <https://doi.org/10.3200/JRLP.140.6.615-618>
- Måttus, R., Allik, J., Realo, A., Rossier, J., Zecca, G., Ah-Kion, J., Amoussou-Y é y é D., Bäckström, M., Barkauskiene, R., Barry, O., Bhowon, U., Björklund, F., Bochaver, A., Bochaver, K., de Bruin, G., Cabrera, H. F., Chen, S. X., Church, A. T., Cissé, D. D., ... Johnson, W. (2012). The Effect of Response Style on Self-Reported Conscientiousness Across 20 Countries. *Personality and Social Psychology Bulletin*, 38(11), 1423–1436. <https://doi.org/10.1177/0146167212451275>
- Peterson, R. A., Rhi-Perez, P., & Albaum, G. (2014). A Cross-National Comparison of Extreme Response Style Measures. *International Journal of Market Research*, 56(1), 89–110. <https://doi.org/10.2501/IJMR-2014-005>
- Rosenberg, M. (1965). Rosenberg self-esteem scale (SES). *Society and the Adolescent Self-Image*.
- Ross, C. E., & Mirowsky, J. (1984). Socially-Desirable Response and Acquiescence in a Cross-Cultural Survey of Mental Health. *Journal of Health and Social Behavior*, 25(2), 189–197. JSTOR. <https://doi.org/10.2307/2136668>
- Singh, J. (1995). Measurement Issues in Cross-National Research. *Journal of International Business Studies*, 26(3), 597–619. <https://doi.org/10.1057/palgrave.jibs.8490188>
- Stening, B. W., & Everett, J. E. (1984). Response Styles in a Cross-Cultural Managerial Study. *The Journal of Social Psychology*, 122(2), 151–156. <https://doi.org/10.1080/00224545.1984.9713475>
- Stigler, J. W., Smith, S., & Mao, L. (1985). The Self-Perception of Competence by Chinese Children. *Child Development*, 56(5), 1259–1270. JSTOR. <https://doi.org/10.2307/1130241>
- Van de Vijver, F. J. R., & He, J. (2014). *Report on Social Desirability, Midpoint and Extreme Responding in TALIS 2013*. <https://doi.org/10.1787/5jxswcft76h-en>
- Van Dijk, T. K., Datema, F., Welten, S., & Van de Vijver, F. J. (2009). *Acquiescence and extremity in cross-national surveys: Domain dependence and country-level correlates*.
- van Herk, H., Poortinga, Y. H., & Verhallen, T. M. M. (2004). Response Styles in Rating Scales: Evidence of Method Bias in Data From Six EU Countries. *Journal of Cross-Cultural Psychology*, 35(3), 346–360. <https://doi.org/10.1177/0022022104264126>
- Wang, R., Hempton, B., Dugan, J. P., & Komives, S. R. (2008). Cultural differences: Why do Asians avoid extreme responses? *Survey Practice*, 1(3).

- Whitcomb, K. M., Önköl, D., Curley, S. P., & Benson, P. G. (1995). Probability judgment accuracy for general knowledge. Cross-national differences and assessment methods. *Journal of Behavioral Decision Making*, 8(1), 51–67. <https://doi.org/10.1002/bdm.3960080105>
- Wright, G. N., & Phillips, L. D. (1980). Cultural variation in probabilistic thinking: Alternative ways of dealing with uncertainty*. *International Journal of Psychology*, 15(1–4), 239–257. <https://doi.org/10.1080/00207598008246995>
- Yates, J. F., & Lee, J. W. (1996). Chinese decision making. *Hong Kong: Oxford University Press*.
- Yates, J. F., Lee, J. W., Levi, K. R., & Curley, S. P. (1990). Measuring and analyzing probability judgment accuracy in medicine. *Philippine Journal of Medicine*, 21–32.
- Yates, J. F., Lee, J.-W., & Bush, J. GG. (1997). General Knowledge Overconfidence: Cross-National Variations, Response Style, and “Reality.” *Organizational Behavior and Human Decision Processes*, 70(2), 87–94. <https://doi.org/10.1006/obhd.1997.2696>
- Yates, J. F., Lee, J.-W., & Shinotsuka, H. (1996). Beliefs about Overconfidence, Including Its Cross-National Variation. *Organizational Behavior and Human Decision Processes*, 65(2), 138–147. <https://doi.org/10.1006/obhd.1996.0012>
- Yates, J. F., Zhu, Y., Ronis, D. L., Wang, D.-F., Shinotsuka, H., & Toda, M. (1989). Probability judgment accuracy: China, Japan, and the United States. *Organizational Behavior and Human Decision Processes*, 43(2), 145–171. [https://doi.org/10.1016/0749-5978\(89\)90048-4](https://doi.org/10.1016/0749-5978(89)90048-4)
- Yu, J. H., Keown, C. F., & Jacobs, L. W. (1993). Attitude Scale Methodology: *Journal of International Consumer Marketing*, 6(2), 45–64. https://doi.org/10.1300/J046v06n02_05
- Zax, M., & Takahashi, S. (1967). Cultural influences on response style: Comparisons of Japanese and American college students. *Journal of Social Psychology*, 3–10.

Table 1. MANOVAs Comparing Response Tendencies: of American and Chinese Managers

	P.R.C.			U.S.			
<i>Extreme Responses-</i>	<u>Mean</u>	<u>S.D.</u>		<u>Mean</u>	<u>S.D.</u>	<u>F-Value</u>	
Factual/informational item	16.2	(5.02)	>	11.6	(7.92)	24.4	**
Nomothetic evaluation	15.2	(4.31)	>	12.6	(7.16)	9.78	*
Idiographic evaluation	1.34	(3.07)	<	5.17	(5.96)	32.87	**
<i>Midpoint Responses</i>							
Factual/informational item	1.23	(1.30)	<	4.86	(4.50)	61.82	**
Nomothetic rating task	1.76	(1.62)	<	3.69	(3.98)	20.77	**
Idiographic rating task	8.01	(6.89)	Na	8.16	(5.18)	.04	
<i>Midrange Responses</i>							
Factual/informational item	5.84	(3.26)	<	11.80	(8.72)	42.27	**
Nomothetic rating task	6.60	(2.84)	<	9.64	(7.62)	14.52	**
Idiographic rating task	23.58	(9.12)	>	20.03	(8.21)	7.62	*

Note. * $p < .01$, ** $p < .001$. MANOVA 1 assessed differences between country groups for the number of extreme (1, 7) and midpoint response (4) scores given by each respondent (Hotelling's test of multivariate significance difference = .54; $p < .001$). Factual memory task involved rating the frequency of usage of various performance criteria; nomothetic task involved rating the importance of criteria, and the idiographic task involved rating the actual performance of the company on these criteria.

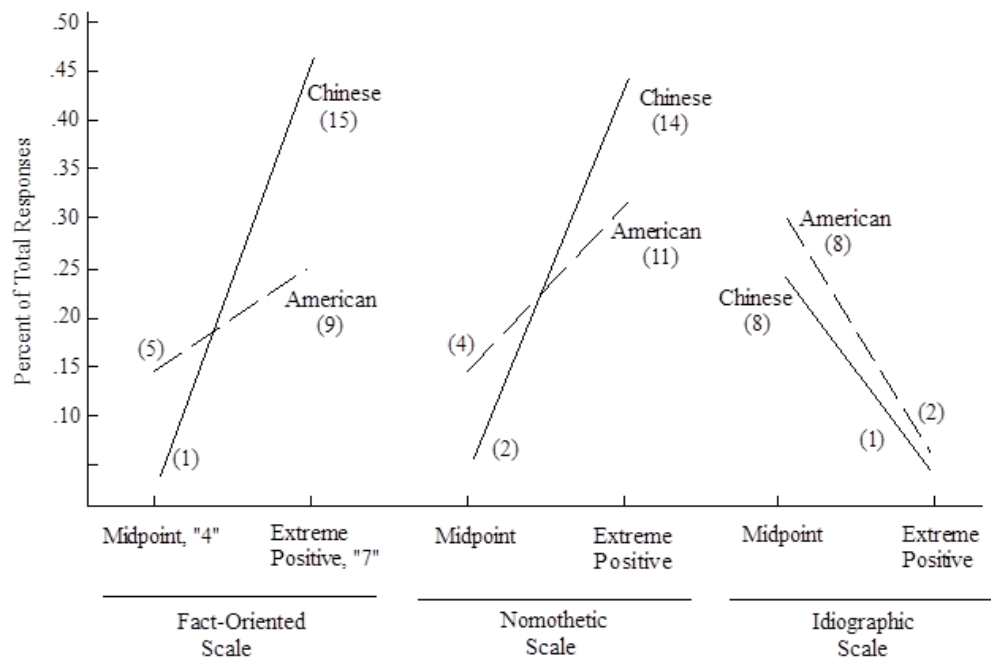


Figure 1. Comparison of Midpoint and Extreme Positive Responses for American and Chinese Managers

Figure 1 shown above are proportions of midpoint ("4") and extreme ("7") responses for respective rating types on the part of Chinese and American managers. Numbers in parentheses indicate the average number of times respondents from each culture group assigned midpoint or extreme scores.

Appendix

Examples of Survey Items and Corresponding Knowledge-Based, Nomothetic, and Idiographic Rating Scales

Item Category (Idiographic)	Item	Frequency (Knowledge-Based)	Importance (Nomothetic)	Actual Performance
		Never Excellent	Regular Minimal	Frequent Moderate Great Poor Okay
<u>Financial</u>	ROI	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7
	ROE			
	Annual profit level			
	Sales growth			
	Ability to control costs			
	Ability to earn foreign exchange			
<u>Marketing</u>	Market share			
	Customer service			
	Customer satisfaction			
	Customer brand awareness			
<u>Operational</u>	Obtaining local materials/parts			
	Obtaining short-term financing			
	Manufacturing quality control			
	Acquiring needed information			
<u>Technology</u>	Technological independence			
	Product design capability			
	Technology development capability			
<u>Relational</u>	JV relations with parent firms			
	Relations between parent firms			
	Obtaining government support			
	Obtaining local contacts			

Note. The explanatory labels, “Knowledge-Based”, “Nomothetic”, and “Idiographic” were not included on the original survey.