



## Cultural differences in response styles: The role of dialectical thinking

Takeshi Hamamura, Steven J. Heine \*, Delroy L. Paulhus

*University of British Columbia, 2136 West Mall, Vancouver, BC, Canada V6T 1Z4*

Received 22 May 2007; received in revised form 15 October 2007; accepted 24 October 2007

Available online 20 December 2007

---

### Abstract

Cultural differences in questionnaire response styles have been reported in a number of studies. Compared to those of European-heritage, responses from individuals of East-Asian heritage tend to be more ambivalent and moderate. These stylistic differences warrant attention because they may contaminate substantive conclusions about cultural differences. One possible source of these stylistic differences is the East-Asian tendency toward *dialectical thinking*, that is, a tolerance of contradictory beliefs. In Study 1 ( $N = 4835$ ), we exemplify the difficulty of demonstrating cultural differences in response styles using the Rosenberg self-esteem scale: the problem is that ambivalence and moderacy indices are inextricably confounded with mean levels of self-esteem. In Study 2 ( $N = 185$ ), we use a broader set of personality measures to show that cultural differences in ambivalence and moderacy can be established independently of mean levels. We go further to show that dialectical thinking mediates the cultural differences in moderacy and ambivalence. Together, these studies demonstrate the importance of considering dialectical thinking when interpreting results from cross-cultural surveys.

© 2007 Elsevier Ltd. All rights reserved.

*Keywords:* Response style; Culture; Dialecticism; Moderacy; Ambivalence

---

---

\* Corresponding author. Tel.: +1 604 822 6908.

E-mail address: [heine@psych.ubc.ca](mailto:heine@psych.ubc.ca) (S.J. Heine).

## 1. Introduction

Cross-cultural researchers continue to rely heavily on self-report questionnaires, largely because of their ease and convenience. Unfortunately, the self-report method is prone to well-known response styles. These include socially desirable responding, acquiescent responding, and extreme responding. A wealth of research on Western respondents indicates that all three can contaminate self-reports (Paulhus & Vazire, 2007).

Accumulating research also indicates that the use of response styles differs across cultural groups. The best documented case concerns *extreme response style* (preference for the highest and lowest response options), which is typically contrasted with a *moderate response style* (preference for middle response options). Compared to North Americans of European-heritage, higher levels of extreme responding have been observed in African-Americans (Bachman & O'Malley, 1984) and Latino Americans (Hui & Triandis, 1989). In contrast, East Asians seem to show more moderacy than samples of European-heritage (Chen, Lee, & Stevenson, 1995).

This burgeoning literature raises concerns about the possible contaminating effect of response styles in cross-cultural research. Some researchers have concluded that response styles such as extremity are not a problem in cross-cultural research (e.g., Chen et al., 1995; Schmitt & Allik, 2005). Others raise concerns that apparent substantive differences across cultures may be artifactual, deriving (at least in part) from differences in response styles (e.g., van de Vijver & Leung, 1997). Debates about how to interpret cultural variation in questionnaire responses would be sharpened if we could confirm that cross-cultural differences in response styles actually exist independently of mean level differences.

### 1.1. Psychological mechanisms responsible for cross-cultural variations in response styles

Cultural tendencies in response styles may simply be superficial response habits that are only manifested in questionnaire behavior. Alternatively, those response style differences may arise from differences in deeper psychological constructs (Smith, 2004): that is, they may reflect the fundamentally different ways that members of different cultures construe themselves and their social worlds. In this paper, we consider a possible role for dialecticism in the manner that people answer questions on Likert scales.

*Dialectical thinking*: Cultures differ in *dialectical thinking*, that is, the tolerance for holding apparently contradictory beliefs (Peng & Nisbett, 1999). In particular, dialectical thinking appears to be more common among East Asians than North Americans. For example, Peng and Nisbett (1999) found that Chinese often endorsed both of two arguments that Americans viewed as incompatible.

Dialectical thinking is also evident in the way that East Asians think of themselves. For example, Choi and Choi (2002) showed that, compared to Americans, Koreans displayed self-beliefs that were less consistent (e.g., I'm outgoing but somewhat shy). Spencer-Rodgers, Peng, Wang, and Hou (2004) found that, compared to Americans, Chinese self-evaluations are more ambivalent in the sense of incorporating both positive and negative self-evaluations.

In the domain of self-reported affect, the relation of positive and negative affect tends to be orthogonal among North Americans but slightly positive among East Asians (Bagozzi, Wong, & Yi, 1999). In sum, a growing body of research indicates that, compared to North Americans, East Asians describe themselves and their feelings in a more dialectical fashion.

The construct of dialecticism may also help us understand cross-cultural variation in questionnaire response styles. First, dialectical thinkers' tolerance of contradiction should reduce the salience of keying direction, that is, the contrast in items designed to measure opposite poles of the same construct. To the extent that they are dialectical thinkers, then, East Asians should see less discrepancy between oppositely-keyed items (Wong, Rindfleisch, & Burroughs, 2003). On an extraversion scale, for example, they should see less contradiction between the true-keyed item "outgoing" and the false-keyed item "shy".<sup>1</sup> Dialectical thinking may thus explain the elevated ambivalence scores observed among East-Asian respondents (Wong et al., 2003).

Second, the greater accessibility of opposing ideas should induce a moderate response style, that is, reduced extremity on items written for either pole. Dialectical thinkers tend to maintain access to contrary views whereas non-dialectical thinkers actively suppress them (Peng & Nisbett, 1999). For instance, a left-leaning dialectical thinker may give moderate responses to both liberal- and conservative-keyed items because contradictory arguments remain available. As dialectical thinkers, then, East Asian respondents should give such moderate responses across the board. In sum, dialectical thinking may explain the known Asian–European differences in both ambivalence and moderate responding.

### *1.2. The present studies*

In two studies, we investigated the role played by dialectical thinking in known cross-cultural differences in two response styles – ambivalent responding and moderate responding. Our cultural focus was on differences between (1) East Asian respondents, (2) North American respondents of East-Asian heritage (AH), and (3) North American respondents of European-heritage (EH). Study 1 investigated the possibility of differentiating response styles from mean evaluative levels in a large cross-cultural archive of responses to the Rosenberg self-esteem scale (RSES). Study 2 extended our investigation to the broader domain of personality and evaluated the possible mediating role of dialectical thinking.

## **2. Study 1: Rosenberg self-esteem scale (RSES)**

Self-esteem scales have drawn much attention from cross-cultural researchers, partly because of research showing the consistently lower scores for East Asians compared to North Americans (e.g., Heine & Hamamura, 2007). These mean differences in self-esteem have been confirmed to operate independently of response styles. Our concern in the present paper is the converse: can response style differences be confirmed independently of mean level differences?

### *2.1. Method*

The data consisted of responses to the Rosenberg self-esteem scale (RSES; Rosenberg, 1965) that were compiled from an archive of cross-cultural studies conducted by Steven Heine and

---

<sup>1</sup> We prefer the terminology of "true-keyed" vs. "false-keyed" over the labels "positive-keyed" vs "negative-keyed" because the latter labels imply desirable and undesirable qualities.

his colleagues.<sup>2</sup> All responses to the 10-item RSES were collected with a 5-point format with three anchor labels: strongly disagree (1), neutral (3), and strongly agree (5).

Participants were categorized into three groups. The European-heritage group (EH) comprised 1478 respondents who self-identified as European-heritage and completed the RSES in North America. The East-Asian heritage group (AH) comprised 1618 respondents who self-identified as Asian-heritage and completed the RSES in North America. A third sample comprised 1739 Japanese college students who completed a Japanese language version of the RSES in Japan.

The three groups were similar in mean age and gender proportions (EH: 62.1% female, mean age 21.1 yrs, AH: 65.9% female, mean age 20.4 yrs, Japanese: 57.3% females, mean age 19.5 yrs). Nonetheless, as a cautionary measure, age and gender were included as covariates in any analysis where they correlated significantly with a dependent variable.

### 2.1.1. *Indices of response style*

*Moderate responding:* Following Chen et al. (1995), moderate responding was measured by counting the number of times out of 10 items that a participant used the middle option on the 5-point scale (i.e., 3). Given past findings, we predicted that the Japanese group would score highest, followed by the AH, and then the EH.

*Ambivalent responding:* A number of statistics have previously been used to index ambivalence: (1) the correlation between true-keyed and false-keyed items (Bagozzi et al., 1999; Wong et al., 2003), (2) absolute differences between true-keyed and false-keyed items (Choi & Choi, 2002), and (3) special ambivalence formulas (Spencer-Rodgers et al., 2004).

We operationalized ambivalence as the absolute difference between true-keyed and false-keyed items. This method directly captures ambivalence for each participant whereas the correlational method can only be calculated on the group. Furthermore, the absolute difference method can be used when the response format includes a neutral option as in our case. Unfortunately, the indices of ambivalence used by Spencer-Rodgers et al. (2004) cannot be used here because our 5-point response format has a midpoint.

To index ambivalence, we calculated the absolute difference between a respondent's ratings of true-keyed and false-keyed items. For example, if a participant's mean rating was 5.0 on true-keyed items and 2.0 on false-keyed items (before scores are reversed), the absolute difference is 3. In essence, this index captures the degree to which the respondent sees the true and false-key items as making opposite claims. For ease of interpretation, the obtained absolute difference was inverted, by subtracting the scores from 5. As a result, a larger number indicates more ambivalent responses. Our Japanese respondents should show the most ambivalence followed by the AH group and then the EH group.

## 2.2. *Results and discussion*

Alpha reliabilities for the RSES were .88, .86, and .76, for the EH, AH, and Japanese samples, respectively. An ANCOVA on mean RSES levels showed significant cultural differences,  $F(2, 4812) = 625.98, p < .001$ . The means fell in the predicted order, that is, 39.6, 35.7, and 31.1 for

<sup>2</sup> Other aspects of these data were reported in a study of the acculturation of self-esteem (Heine & Lehman, 2004).

the EH, AH, and Japanese samples, respectively. Similar cultural differences in level of self-esteem have been found in much prior work (for reviews see Heine & Hamamura, 2007; Heine, Lehman, Markus, & Kitayama, 1999).

### 2.2.1. Response styles

*Moderate responding:* The ANCOVA was also significant for cultural differences in moderate responding,  $F(2, 4813) = 306.59, p < .001$ . Again, the mean values fell in the predicted direction. Japanese respondents ( $M = 2.70, SD = 1.92$ ) chose the midpoint of the scale more frequently than EH respondents ( $M = 1.15, SD = 1.30$ ) and AH respondents ( $M = 1.96, SD = 1.79$ ). These differences were pronounced in magnitude ( $d = .95$  between Japanese and EH,  $d = .40$  between Japanese and AH, and  $d = .52$  between the two North American samples). These values are similar to those found by Chen and Colleagues (1995).

*Ambivalent responding:* The ANCOVA was also significant for the ambivalence index,  $F(2, 4812) = 403.14, p < .001$ . As predicted, Japanese ( $M = 3.95, SD = .83$ ) provided more ambivalent responses compared to AH ( $M = 3.48, SD = .97$ ) and EH respondents ( $M = 2.98, SD = 1.03$ ). These differences were also pronounced in magnitude with effect sizes of  $d = 1.05$  between Japanese and EH,  $d = .52$  between Japanese and AH, and  $d = .50$  between the two North American samples. These findings replicate those of Spencer-Rodgers and colleagues (2004).

### 2.2.2. Relations between scale means and response styles

Missing from previous reports is the pattern of associations among mean levels and response styles. The intercorrelations and reliabilities of the measures (Cronbach's alpha for self-esteem scale and the Spearman–Brown coefficients for moderacy and ambivalence) are reported in Table 1. Note that the cultural differences in means and response styles noted above are also reflected in the significant correlations in the final column.

Ambivalence was highly correlated with moderacy ( $r = .59$ ). This overlap results at least partly from the fact that a preference for midpoints necessarily increases ambivalence scores. Hence it is important to confirm the effects of ambivalence after controlling for moderacy (but not the reverse as there is no reason to suspect that ambivalence would influence moderacy). For example, the correlation of ethnicity with ambivalence remained significant after partialing out moderacy ( $r = .24, p < .001$ ). This analysis assures us that indices of ambivalence and moderacy are not redundant.

Table 1  
Correlations between RSES means and two response styles (Study 1)

	Self-esteem	Moderate responding	Ambivalent responding	East-Asian ethnicity <sup>a</sup>
Self-esteem	(.86)	-.41	-.65	-.46
Moderate responding		(.47)	.59	.34
Ambivalent responding			(.72)	.38
East-Asian ethnicity <sup>a</sup>				(1.00)

Note.  $N = 4835$ .

All correlations are significant at the .01 level, two-tailed.

Values in the diagonal are reliabilities.

<sup>a</sup> Scored as European-heritage (EH) = 1; East-Asian heritage (AH) = 2; Japanese = 3.

Both ambivalence and moderacy were negatively associated with RSES means ( $r = -.65$  and  $r = -.41$ , respectively): that is, those with high self-esteem tended to respond in an un-ambivalent and extreme fashion to the items. Similar associations were found within all three cultural groups.

These associations raise the possibility that observed cultural differences in response styles are merely an artifact of differences in scale means. In other words, most of those with relatively low RSES scores scored around 30 – the midpoint of the scale. Hence low RSES scores could result from having relatively few positive thoughts about oneself or from having an ambivalent or moderate response style. In short, self-esteem scores and response styles are confounded here.

In sum, Study 1 confirmed that individuals with greater exposure to East Asian culture exhibit more moderate and more ambivalent responses compared to individuals with greater exposure to Western culture. Unfortunately, the study also raised the fear that these response style differences derive with cultural differences in scale means. We attempt to eliminate this confound in Study 2.

### 3. Study 2: personality data

To circumvent the limitations of Study 1, we turned to a domain that was (a) broader in scope, and (b) more substantive than evaluative. The domain of personality traits satisfies these criteria. Factor analyses of item sets designed to cover the full personality space typically yield five super factors known as the “Big Five” (e.g., Costa & McCrae, 1989; Goldberg, 1992; John & Srivastava, 1999). Rather than evaluation, these factors are held to capture the substance of personality. By using items from all five factors, we also hoped to avoid dimension-specific artifacts. We also explored the mediating role of dialectical thinking.

#### 3.1. Method

We collected data from a total of 185 student volunteers (65% female) at the University of British Columbia. The sample was partitioned into two cultural groups based on language spoken at home: the European-heritage (EH) group spoke a European language at home ( $n = 72$ ) and the East-Asian heritage (AH) group spoke an East-Asian language at home ( $n = 86$ ). The 27 students who spoke other languages or multiple languages at home were excluded from the data analyses. Participants received extra credit toward their course grade.

The questionnaire included the dialectical self scale (Spencer-Rodgers, Srivastava, & Peng, 2001). Responses to the dialectical self scale were indicated in a 7-point format with a neutral option. Sample items include “When I hear two sides of an argument, I often agree with both,” and “I sometimes believe two things that contradict each other”.

In addition, we included a Big Five inventory consisting of 60 adjectives (Wada, 1996). Designed for use in both English and Japanese, the items were derived from Goldberg’s (1992) Big Five markers and compared against the Big Five Inventory (John & Srivastava, 1999). We added 15 extra items written to better balance the number of true- and false-keyed items on each personality factor because of our special interest in keying issues. A 7-point response format with a neutral option in the middle was used for all the items. Finally, participants were asked to report their age, gender, and high school grade point average (GPA).

### 3.2. Results and discussion

#### 3.2.1. Comparability of samples

On average, the EH participants ( $M = 21.46$ ) were more than a year older than the AH participants ( $M = 20.28$ ),  $t(156) = 2.40$ ,  $p < .05$ . Moreover, age was significantly correlated with dialecticism: hence age was entered as a covariate for that analysis. The two groups did not differ on gender proportion,  $\chi^2(1, N = 155) < 1$ , *ns*, hence gender is mentioned no further. Also, the two groups did not differ significantly on high school GPA,  $t(152) < 1$ , *ns*.

#### 3.2.2. Personality composite scores

Of the 75 items in the personality scale, the EH and AH participants differed significantly in their mean responses for 49 items. To avoid the confound between means and response styles, we focused on the 26 items that did not differ between cultures. They included coverage of all of the Big Five factors. Subsequent analyses were calculated only on this reduced item set.

#### 3.2.3. Response styles

As in Study 1, moderacy was calculated as the number of times (out of 26 items) that a participant endorsed the midpoint. Again, ambivalence was operationalized as the absolute difference between responses to true-keyed items and false-keyed items (before reversing the false-keyed items). The mean of these difference scores was then reversed by subtracting them from 7. After this transformation, high scores can be directly interpreted as ambivalence.

*Moderate responding:* Replicating Study 1, AH respondents were more likely to select the scale midpoint ( $M = 4.91$ ,  $SD = 3.41$ ) than were EH respondents ( $M = 3.94$ ,  $SD = 2.67$ ),  $t(156) = 1.95$ ,  $p = .05$ . The effect size ( $d = .32$ ) was slightly smaller than that observed in Study 1 ( $d = .52$ ).

*Ambivalence:* Compared to the EH respondents ( $M = 5.51$ ,  $SD = .77$ ), AH respondents ( $M = 5.79$ ,  $SD = .78$ ) provided more ambivalent responses,  $t(156) = 2.26$ ,  $p < .05$ . Again this cultural difference ( $d = .36$ ) was slightly smaller than that observed in Study 1 ( $d = .50$ ).

Note again that the item sets we used here did not show any cultural difference in mean levels. Hence the observed cultural differences in these two responses styles cannot be explained away as an artifact of the mean differences – as they could in Study 1.

#### 3.2.4. Dialectical self scale

Alpha reliabilities for the dialectical self scale were .79 and .83, for EH and AH respondents, respectively. Consistent with past research, AH respondents ( $M = 4.11$ ,  $SD = .54$ ) were significantly more dialectical than EH respondents ( $M = 3.87$ ,  $SD = .51$ ),  $F(1, 155) = 7.51$ ,  $p < .01$ ,  $d = .46$ .

#### 3.2.5. Relations between dialectical thinking and response styles

Table 2 provides the correlations among culture, dialectical thinking, moderacy, and ambivalence. Note that dialectical thinking was significantly correlated with both ambivalence ( $r = .29$ ,  $p < .001$ ) and moderacy ( $r = .22$ ,  $p < .01$ ). As in Study 1, the two response styles overlapped considerably ( $r = .40$ ,  $p < .001$ ). Nonetheless, the association of ambivalence with dialectical thinking

Table 2  
Correlations among ethnicity, dialectical thinking and two response styles (Study 2)

	Dialectical thinking	Moderate responding	Ambivalent responding	East-Asian ethnicity <sup>a</sup>
Dialectical thinking	(.82)	.21	.27	.22
Moderate responding		(.62)	.40	.15
Ambivalent responding			(.68)	.18
East-Asian ethnicity <sup>a</sup>				(1.00)

Note.  $N = 185$ .

All correlations are significant at the .05 level, two-tailed.

Values in the diagonal are reliabilities.

<sup>a</sup> Scored as European-heritage (EH) = 1; East-Asian (AH) heritage = 2.

remained significant after controlling for moderacy ( $r = .21, p < .01$ ). As in Study 1, ethnicity was also significantly associated with both ambivalence and moderacy.<sup>3</sup>

Next, mediation analyses were conducted to see if dialecticism mediated the cultural differences in response styles (Baron & Kenny, 1986). A multiple regression analysis in which ambivalence was predicted simultaneously by culture and dialecticism revealed that dialecticism not only remains as a significant predictor ( $\beta = .24, p < .01$ ) but eliminates the role of culture as a predictor of ambivalence: in fact, culture was no longer a significant predictor of ambivalence ( $\beta = .13, p > .10$ ). A test of indirect effects (i.e., the impact of culture on response style via dialectical thinking) was conducted by the bootstrap procedure (Shrout & Bolger, 2002). Five thousand bootstrap resamples were obtained. The indirect effect of culture on ambivalence via dialecticism was significant,  $ab = .05$ ,  $SE = .03$ ,  $CI_{95} = .01, .12$ . Hence, dialecticism mediates the cultural difference in ambivalence.

Similarly, when dialecticism was entered along with culture in predicting moderate responding, dialecticism remained a significant predictor ( $\beta = .18, p < .05$ ) but culture was no longer significant ( $\beta = .12, p > .15$ ). The indirect effect of culture on moderacy via dialecticism was significant,  $ab = .04$ ,  $SE = .02$ ,  $CI_{95} = .01, .09$ . Hence, dialecticism also mediated the cultural difference in moderate responding.

One might argue the converse – that dialectical thinking is mediated by response styles. Theoretically speaking, dialectical thinking is held to be a more fundamental psychological construct with other testable correlates. Response styles are usually assumed to operate only on questionnaire behavior. The observed pattern of results also disputes the reverse ordering: specifically, the mean of the dialectical self scale sits almost right on the midpoint of four for both EH and AH: hence, neither response style could elevate observed scores on the dialectical thinking scale.

#### 4. General discussion

Some writers remain concerned about the impact of response styles on the validity of cross-cultural survey research (e.g., van de Vijver & Leung, 1997; Wong et al., 2003) whereas others dismiss

<sup>3</sup> After controlling for moderacy, the correlation between ethnicity and ambivalence dropped below significance ( $r = .11, p > .10$ ). Rather than a contradiction with the  $r = .24$  in Study 1, this lower value in Study 2 results from the restricted range of ethnicity scores (2 vs. 3 cultural groups).



that concern (e.g., [Chen et al., 1995](#)). Our goal was to confirm the alleged cultural differences in response styles and investigate dialectical thinking as a possible mechanism. We confirmed that there are genuine cultural differences in response styles: compared to North Americans, participants of East-Asian heritage exhibit more ambivalent and moderate responding on self-report items. This difference creates difficulty for the interpretation of mean scores across cultures: it is not clear whether there are substantive cultural differences or just cultural differences in these response styles.

Although some of these response style differences have been found before (e.g., [Chen et al., 1995](#); [Wong et al., 2003](#)), those studies did not incorporate controls for mean cultural differences. As we showed in Study 1, mean differences in self-esteem are inextricably confounded with these two response styles. Therefore, the Rosenberg Self-esteem scale, despite its popularity, is not appropriate for the study of cultural differences in response styles.

In principle, this artifact extends to all self-evaluation measures in the literature because their item means tend to fall well above the midpoint ([Paulhus & Vazire, 2007](#)). To report that one has an especially positive self-evaluation requires the endorsement of extreme options. In light of this confound, our goal of confirming cultural differences in response styles required a broader set of self-report dimensions. The Big Five personality questionnaire provided items from five important and orthogonal substantive domains. This choice furnished enough items to form a set of 26 items having no mean differences between cultures. With these controls in place, we were able to confirm that respondents with East-Asian heritage still showed more moderation and ambivalence in their answers to survey items.

We also pursued the possibility that these response styles – typically viewed as relatively superficial questionnaire behaviors – have their roots in deeper psychological proclivities. In particular, we hypothesized that the East Asian tendency toward dialectical thinking might be responsible for the observed cultural differences in both moderation and ambivalence. Our Study 2 supported this notion in that dialecticism completely mediated the cultural differences in both response styles. People who are more likely to tolerate contradiction are also more likely to (a) endorse the neutral response option and (b) endorse apparently contradictory statements.

Finally, in both studies, we found substantial intercorrelations between ambivalent and moderate responding. The sizable intercorrelations and the fact that dialecticism mediated cultural differences in both styles suggest that they are manifestations of the same underlying psychological process. That is, East Asians exhibit more moderation and ambivalence on questionnaires because they have a more dialectical way of thinking about themselves.

#### *4.1. Practical implications and future research*

Papers on response style often raise the issue of whether stylistic effects should be controlled or not (e.g., [Paulhus, 1986](#)). To the extent that stylistic differences reflect cultural modes of thought, then controlling for their effects may not be appropriate ([Smith, 2004](#)). Thoughtless partialing of response styles may result in the throwing out of the baby with the bathwater. If possible, efforts are best expended in precluding response styles rather than attempting to undo their effects post hoc ([Paulhus, 1991](#)).

We caution that our results may have hinged on our choices for operationalizing response styles. For example, scores on our index of ambivalence are necessarily elevated by high scores

on our measure of moderacy. Although we were able to confirm the independent impact of ambivalent responding, future research should utilize non-overlapping indices. The ideal approach would attempt to replicate our findings using several assessment methods in a multiple-method approach. The inclusion of the ambivalence measures developed by [Spencer-Rodgers and colleagues \(2004\)](#) is one obvious improvement. Another is the addition of responses collected in a forced-choice format: this would preclude moderate responding while permitting the operation of ambivalence. Content coding of ambivalence and moderacy in essays might confirm whether dialectical thinking is manifested in the free-description mode of expression.

We encourage others to join our inquiry into the psychological processes underlying response styles (e.g., [Krosnick, 1999](#)) because – despite recommendations to the contrary (e.g., [Heine, Lehman, Peng, & Greenholtz, 2002](#)) – cross-cultural research is unlikely to shift away from survey methods. For the field to progress with self-report methodology, it is critical that we remain cautious about accepting at face value the cultural differences inferred via survey alone.

## References

- 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*, 491–509.
- Bagozzi, R. P., Wong, N., & Yi, Y. (1999). The role of culture and gender in the relationship between positive and negative affect. *Cognition and Emotion*, *13*, 641–672.
- Baron, R. M., & Kenny, D. (1986). The moderator mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173–1182.
- 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.
- Choi, I., & Choi, Y. (2002). Culture and self-concept flexibility. *Personality and Social Psychology Bulletin*, *28*, 1508–1517.
- Costa, P. T., & McCrae, R. R. (1989). *Manual for the NEO Personality Inventory: Five Factor Inventory/NEO-FFI*. Odessa, FL: PAR.
- Goldberg, L. R. (1992). The development of markers for the Big Five structure. *Psychological Assessment*, *4*, 26–42.
- Heine, S. J., & Hamamura, T. (2007). In search of East Asian self-enhancement. *Personality and Social Psychology Review*, *11*, 1–24.
- Heine, S. J., & Lehman, D. R. (2004). Move the body, change the self: Acculturative effects on the self-concept. In M. Schaller & C. Crandall (Eds.), *Psychological foundations of culture* (pp. 305–331). Mahwah, NJ: Erlbaum.
- Heine, S. J., Lehman, D. R., Markus, H. R., & Kitayama, S. (1999). Is there a universal need for positive self-regard? *Psychological Review*, *106*, 766–794.
- Heine, S. J., Lehman, D. R., Peng, K., & Greenholtz, J. (2002). What's wrong with cross-cultural comparisons of subjective Likert scales? The reference-group problem. *Journal of Personality and Social Psychology*, *82*, 903–918.
- Hui, C. H., & Triandis, H. C. (1989). Effects of culture and response format on extreme response style. *Journal of Cross-Cultural Psychology*, *20*, 296–309.
- John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 102–139). New York: Guilford Press.
- Krosnick, J. A. (1999). Survey research. *Annual Review of Psychology*, *50*, 537–567.
- Paulhus, D. L. (1986). Self-deception and impression management in test responses. In A. Angleiter & J. S. Wiggins (Eds.), *Personality assessment via questionnaire: Current issues in theory and measurement*. Berlin: Springer-Verlag.
- Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17–59). New York: Academic Press.
- Paulhus, D. L., & Vazire, S. (2007). The self-report method of personality assessment. In R. W. Robins, R. C. Fraley, & R. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 224–239). New York: Guilford.

- Peng, K., & Nisbett, R. E. (1999). Culture, dialectics, and reasoning about contradiction. *American Psychologist*, *54*, 741–754.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- Schmitt, D. P., & Allik, J. (2005). Simultaneous administration of the Rosenberg self-esteem scale in 53 nations: Exploring the universal and culture-specific features of global self-esteem. *Journal of Personality and Social Psychology*, *89*, 623–642.
- Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, *7*, 422–445.
- Smith, P. B. (2004). Acquiescent response bias as an aspect of cultural communication style. *Journal of Cross-Cultural Psychology*, *35*, 50–61.
- Spencer-Rodgers, J., Srivastava, S., & Peng, K. (2001). *The dialectical self scale*. Berkeley: University of California. Unpublished data.
- Spencer-Rodgers, J., Peng, K., Wang, L., & Hou, Y. (2004). Dialectical self-esteem and East–West differences in psychological well-being. *Personality and Social Psychology Bulletin*, *30*, 1416–1432.
- van de Vijver, F. J. R., & Leung, K. (1997). *Methods and data analysis for cross cultural research*. Thousand Oaks, CA: Sage Publications.
- Wada, S. (1996). Construction of the Big Five scales of personality trait terms and concurrent validity with NPI. *Japanese Journal of Psychology*, *67*, 61–67.
- 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.