

Inclusion of theory-relevant moderators yield the same conclusions as Sedikides, Gaertner, and Vevea (2005): A meta-analytical reply to Heine, Kitayama, and Hamamura (2007)

Constantine Sedikides,¹ Lowell Gaertner² and Jack L. Vevea³

¹University of Southampton, Southampton, UK; and ²University of Tennessee, Knoxville, Tennessee and

³University of California, Santa Cruz, California, USA

Heine, Kitayama and Hamamura (2007) attributed the Sedikides, Gaertner and Vevea (2005) findings to the exclusion of six papers. We report a meta-analysis that includes those six papers. The Heine *et al.* conclusions are faulty, because of a misspecified meta-analysis that failed to consider two moderators central to the theory. First, some of their effect sizes originated from studies that did not empirically validate comparison dimensions. Inclusion of this moderator evidences pancultural self-enhancement: Westerners enhance more strongly on individualistic dimensions, Easterners on collectivistic dimensions. Second, some of their effect sizes were irrelevant to whether enhancement is correlated with dimension importance. Inclusion of this moderator evidences pancultural self-enhancement: Both Westerners and Easterners enhance on personally important dimensions. The Sedikides *et al.* conclusions are valid: Tactical self-enhancement is pancultural.

Key words: pancultural self-enhancement, self-enhancement, tactical self-enhancement.

Introduction

Heine, Lehman, Markus, and Kitayama (1999) concluded that self-esteem is culturally specific. In their own words, ‘The empirical literature provides scant evidence for a need for positive self-regard among Japanese’, ‘the need for self-regard must be culturally variant’, and ‘the need for self-regard . . . is not a universal, but rather is rooted in significant aspects of North American culture’ (p. 766). Easterners (East Asians, and Japanese in particular) not only lack high self-esteem, but they do not even have a need for self-esteem. Only Westerners (e.g. Americans, Canadians, Northern Europeans) have high self-esteem and a need for it. By implication, only Westerners self-enhance.

We (Sedikides, Gaertner, & Toguchi, 2003) argued that their claim was implausible, given the evidence: (i) in favour of the genetic underpinnings (Neiss *et al.*, 2005) and existential relevance (Pyszczynski, Greenberg, Solomon, Arndt, & Schimel, 2004) of self-esteem; (ii) that positive implicit self-esteem is universal (Yamaguchi *et al.* in press); and (iii) that self-enhancement is a marker of good mental health both in the West (Taylor, Lerner, Sherman, Sage, & McDowell, 2003) and the East (Kobayashi & Brown, 2003). Instead, we offered an alternative hypothesis.

We hypothesized that both Easterners and Westerners self-enhance on domains that are personally important. It so happens that these domains are different for the two cultural groups. The collectivistic domain is more personally important to Easterners, the individualistic domain to Westerners. Hence, Easterners will self-enhance on collectivistic domains, Westerners on individualistic domains. Self-enhancement is a universal esteem-seeking strategy. Our hypothesis was supported in two primary experiments (Sedikides *et al.*, 2003). Challenged by Heine (2005), we accumulated further support for our hypothesis in a meta-analysis (Sedikides, Gaertner, & Vevea, 2005). Heine, Kitayama, and Hamamura (2007) provided another challenge to our hypothesis.

Clarifying the exclusion criteria

In a previously published paper, Heine *et al.* (2007) argued that Sedikides *et al.* (2005) did not include ‘six previously published papers that also tested the Hypothesis but yielded findings directly counter to the claims of Sedikides *et al.* (2003)’. They further stated that ‘we are unclear as to the rationale of Sedikides *et al.* for excluding these studies, as the only mention of these studies in their 2005 paper is in a footnote on page 540 which reads “These criteria identify a subset of studies that are relevant to our framing of the research question. There are other studies on this general topic that are not included, such as Heine and Lehman (1995), Heine *et al.* (2001), and Kitayama *et al.* (1997)”’ (Heine *et al.*, 2007).

Correspondence: Constantine Sedikides, Center for Research on Self and Identity, School of Psychology, University of Southampton, Southampton SO17 1BJ, England, UK. Email: cs2@soton.ac.uk

Received 4 October 2006; accepted 21 February 2007.

We believe that we (Sedikides *et al.*, 2005) mentioned the broad exclusion grounds in our p. 540 footnote, and we also detailed exclusion reasons for Heine and Renshaw (2002) in our footnote 6 on p. 546. Regardless, we submit that a meta-analysis is under scholarly constraints to: (i) propose a theoretical formulation; (ii) set theory-testing parameters by defining the domain of empirical inquiry; and (iii) select studies that fit that domain (Johnson & Eagly, 2000). We followed these principles in our meta-analysis (Sedikides *et al.*, 2005), and this resulted in our exclusion criteria. What were these criteria? They were as follows (Sedikides *et al.*, 2005, p. 540):

First, studies had to sample either members of Western or Eastern culture. Such studies identified cultural membership on the basis of nationality (e.g. Japan, United States), ethnicity (e.g. Asian), or self-construal (Singelis, 1994). Second, studies had to provide a measure of one's perception of self relative to others. Participants in such studies (a) performed a distribution task in which they estimated the percentage of the population from which they are better in terms of a given attribute (b) rated separately self and other . . . , or (c) rated simultaneously self and other on a bipolar scale anchored at the extremes by self and other. Finally, studies had to assess explicitly the self-other comparison on attributes associated with individualism versus collectivism. Participants in such studies compared self and other in regard to individualistic traits or behaviors or collectivistic traits or behaviors.

We provide a brief rationale for the exclusion of each paper in Appendix I.

Our new meta-analysis

The purpose of the rationale is not merely retrospective. Instead, the rationale is relevant to a new meta-analysis that we undertook as a response to Heine *et al.* (2007). In the present article, we heeded Heine *et al.*'s call to expand the scope or criteria of the meta-analysis by including the six papers. We incorporated two of these papers (Markus & Kitayama, 1991;¹ Heine & Lehman, 1995²) in Investigation 1, and the remaining four papers (Kitayama, Markus, Matsumoto, & Norasakkunkit, 1997; Heine & Lehman, 1999; Heine *et al.*, 2001; Heine & Renshaw, 2002) in Investigation 2.

More to the point, the empirical viability of our hypothesis rests on two theory-relevant moderators, which Heine *et al.* (2007) did not include in their meta-analysis. The first moderator is the empirical validation of the comparison dimension (i.e. individualistic *vs* collectivistic). Our theoretical formulation predicts that pancultural self-enhancement will be manifested on dimensions that are empirically validated as either individualistic (i.e. Western self-enhancement) or collectivistic (i.e. Eastern self-

enhancement). Unvalidated dimensions are non-diagnostic and, thus, incapable of providing a test of the theory. Unfortunately, Heine *et al.* (2007) ignored this distinction and aggregated across effects from validated and unvalidated studies. The second moderator is personal importance of the dimension (i.e. important *vs* unimportant). Our theoretical formulation predicts that pancultural self-enhancement will be manifested by both Westerners and Easterners on dimensions of high (rather than low) personal importance. Heine *et al.* (2007) obscured assessment of this possibility by considering what we regard as irrelevant effects from our perspective. Omitting the validation issue and including irrelevant effects results in misspecified meta-analyses, which leads to two inevitable, if not grave, consequences: inappropriate theory-testing and misleading conclusions.

Investigation 1

This investigation addressed Heine *et al.*'s (2007) failure to take into consideration that some effect sizes originated from studies which did not empirically validate the comparison dimensions. Our investigation includes this vital and theory-relevant moderator: samples varied on whether the researchers empirically validated the individualistic and collectivistic dimensions of comparison. Is pancultural self-enhancement observed on validated (rather than unvalidated) comparison dimensions, as Sedikides *et al.* (2005; Investigation 1) proposed and found?

Method

Data sources. Data were obtained from: (i) the eight studies included in Investigation 1 of Sedikides *et al.*'s (2005) meta-analysis, which yielded 15 effect sizes from 15 samples; (ii) the three additional studies included in Investigation 1 of Heine *et al.*'s meta-analysis (2007), which yielded eight effect sizes from six samples; and (iii) a recently completed study (Gaertner, Sedikides, & Chang, 2006³), which yielded one effect size from one sample. For convenience, we refer to these three sources of the 24 effect sizes as Wave 1, 2, and 3, respectively.

Participants in each sample rated self and other on individualistic and collectivistic comparison dimensions. As displayed in Table 1, each sample represented either Eastern or Western culture. Also, samples varied as to whether the researchers empirically validated the individualistic and collectivistic dimensions.

Calculation of effect sizes. The effect of interest examines whether individuals more favourably compare self with other on individualistic versus collectivistic comparison dimensions. Following Sedikides *et al.* (2005), we calcu-

Table 1 Characteristics of the 24 effect sizes included in Investigation 1

Wave	Study	Culture	Validate	Effect size	Standard error
1	Heine and Lehman (1997), Study 1	W	N	-0.4772	0.1656
1	Heine and Lehman (1997), Study 1	E	N	-0.2449	0.1587
1	Hornsey and Jetten (2005), Study 1	W	Y	0.9550	0.2306
1	Hornsey and Jetten (2005), Study 2	W	Y	0.3602	0.3566
1	Hornsey and Jetten (2005), Study 2	E	Y	-0.6825	0.4044
1	Kurman (2001), Study 1	W	Y	0.0312	0.1245
1	Kurman (2001), Study 1	E	Y	-0.1842	0.1185
1	Kurman (2001), Study 2	W	Y	0.1714	0.1185
1	Kurman (2001), Study 2	E	Y	-0.5946	0.1366
1	Norasakkunkit and Kalick (2002)	W	N	-0.8633	0.1287
1	Norasakkunkit and Kalick (2002)	E	N	-0.3823	0.1185
1	Sedikides <i>et al.</i> (2003), Study 1	W	Y	0.7024	0.2306
1	Sedikides <i>et al.</i> (2003), Study 1	E	Y	-0.6176	0.2290
1	Sedikides <i>et al.</i> (2003), Study 2	W	Y	1.3317	0.2261
1	Sedikides <i>et al.</i> (2003), Study 2	E	Y	-1.7605	0.2412
2	Heine and Lehman (1995), Study 2	W	N	-0.4700	0.1448
2	Heine and Lehman (1995), Study 2	W	N	0.4100	0.1459
2	Heine and Lehman (1995), Study 2	E	N	0.3300	0.1417
2	Heine and Lehman (1995), Study 2	E	N	0.8700	0.1480
2	Markus and Kitayama (1991)	W	N	0.4170	0.1516
2	Markus and Kitayama (1991)	E	N	0.3582	0.1511
2	Ross <i>et al.</i> (2005), Study 1	W	N	-0.1510	0.1484
2	Ross <i>et al.</i> (2005), Study 1	E	N	-0.2930	0.1355
3	Gaertner <i>et al.</i> (2006)	E	Y	-0.3024	0.1836

E, Easterners; N, no; W, Westerners; Y, yes.

lated an effect size by: (i) subtracting the mean self-other rating on the collectivistic dimension from the mean self-other rating on the individualistic dimension; and (ii) standardizing that difference by the sample standard deviation of the within-subject difference score. The standard deviation of the difference score is appropriate, because the effect of interest conceptually is a within-subject effect and each sample assessed that effect with a within-subject method.

A positive effect size indicates that the self-other comparison was more favourable to self on the individualistic dimension. A negative effect size indicates that the self-other comparison was more favourable to self on the collectivistic dimension. An effect size of zero indicates that the self-other comparison was equally favourable to self on the individualistic and collectivistic dimensions.

Results and discussion

We created 12 data sets to address three issues. The first issue was raised by Heine *et al.* (2007). They were concerned that the Eastern and Western effect sizes from Sedikides *et al.* (2003; Study 2) biased the results, because participants were Americans who were identified as having either an independent or interdependent self-construal.

Consequently, half of our data sets included those effect sizes, whereas the other half excluded them.

The second issue was introduced by the studies that Heine *et al.* (2007) included in their meta-analysis. As displayed in Wave 2 of Table 1, the two Western effects sizes and the two Eastern effect sizes, respectively, from Heine and Lehman (1995) were derived from the same samples of participants. Inclusion of these samples in the same analysis violates the independence assumption. We satisfied that assumption with data sets that varied which non-independent samples were excluded. Half of the data sets (labelled Set 1) excluded the second Western and Eastern effect size from Heine and Lehman (1995; i.e. effect sizes = 0.41 and 0.87), whereas the other half (labelled Set 2) excluded the first Western and Eastern effect size from that study (i.e. effect sizes = -0.47 and 0.33).

The third problem was introduced by the standardizing agent of the effect size; that is, the standard deviation of the difference score (Heine *et al.*, 2007). Although that standardizing agent reflects the within-subject nature of self-enhancement (i.e. subjective perception of self vs other), the standard deviation of the difference score overestimates the effect size to the extent to which the correlation between the components of the difference score (i.e. self-enhancement on independent vs interdependent dimen-

Table 2 Self-other comparisons on individualistic versus collectivistic dimensions as a function of culture and validate

Data set	Validate Yes				Validate No				Culture × Validate	
	Western	Eastern	<i>Q</i>	<i>p</i> =	Western	Eastern	<i>Q</i>	<i>p</i> =	<i>Q</i>	<i>p</i> =
Including Sedikides <i>et al.</i> (2003; Study 2)										
1	0.57*	-0.66*	21.95	0.0000	-0.31	-0.05	0.96	0.3271	15.81	0.0001
2	0.57*	-0.67*	17.92	0.0000	-0.14	0.06	0.40	0.5271	11.54	0.0007
Excluding Sedikides <i>et al.</i> (2003; Study 2)										
1	0.40*	-0.44*	13.45	0.0002	-0.31*	-0.05	1.56	0.2117	12.59	0.0004
2	0.42*	-0.45*	9.97	0.0016	-0.14	0.06	0.56	0.4542	7.95	0.0048

* $p < 0.05$. *Q*-tests for the Culture × Validate effect and the culture effect in levels of validate are chi-squared on 1 degree of freedom. A positive effect size indicates a more self-favouring social comparison on the individualistic than on the collectivistic dimension. A negative effect size indicates a more self-favouring social comparison on the collectivistic than on the individualistic dimension.

sions) decreases (Morris & DeShon, 2002; Becker, 2006). To rule out the possibility that any conclusions vary as a function of a potential correlation, we transformed Data Sets 1 and 2, with and without Sedikides *et al.* (2003; Study 2), to reflect: (i) no correlation (i.e. $\rho = 0.00$); (ii) a moderate correlation (i.e. $\rho = 0.50$ – this is equivalent to the untransformed effect sizes as they are presented in Table 1); and (iii) a strong correlation ($\rho = 0.75$). The transformation involves calculating an effect size in which the squared standard-deviation-of-the-difference-score is divided by: $2 - 2 * \rho$. Stated somewhat differently, we performed the subsequent analysis on Data Sets 1 and 2, with and without Sedikides *et al.* (2003; Study 2) transformed to reflect $\rho = 0.00, 0.50$, and 0.75 . For expediency we present results for $\rho = 0.50$. Assuredly, analyses for $\rho = 0.00$ and 0.75 yield the same conclusions as do those for $\rho = 0.50$.

We planned to approach the analysis with a mixed-effects model, because: (i) it enables a broader context of generalization; and (ii) data typically violate the homogeneity assumption of the fixed-effects model (National Research Council, 1992; Hedges & Vevea, 1998; Hunter & Schmidt, 2000; Field, 2001, 2003). We regressed effect sizes on culture, validate and the Culture–Validate interaction, and proceeded to estimate parameters with unconditional maximum likelihood using a program developed for that purpose (Veeva & Hedges, 1995; Veeva & Woods, 2005). Each analysis evidenced a significant *Q*-test of the fixed-effects model's homogeneity assumption, which bolsters our choice for the mixed-effects model.

As displayed in Table 2, all data sets yielded the same conclusion. Consistently with Sedikides *et al.* (2005; Investigation 1), each analysis evidenced a Culture–Validate interaction. For studies that validated the comparison dimensions, there was a significant effect of culture. Westerns evidenced a more self-favouring social comparison on the individualistic than collectivistic dimension, whereas Easterners evidenced a more self-

favouring social comparison on the collectivistic than on the individualistic dimension. For studies that did not validate the comparison dimensions, however, there was no effect of culture. Both Westerners and Easterners evidenced social comparisons that were no more self-favouring on what was ostensibly an individualistic than a collectivistic dimension of comparison. The methodologically superior studies substantiate the argument that both Westerners and Easterners strategically self-enhance on comparison dimensions that are uniquely infused by their culture.

Investigation 2

Although Western and Eastern cultures self-enhance on different comparison dimensions (i.e. individualistic vs collectivistic), the pancultural self-enhancement formulation anticipates that the same process underlies the differential expression: people self-enhance on personally important comparison dimensions. Such was the pattern reported by Sedikides *et al.* (2005; Investigation 2). Will the present Investigation 2 replicate this pattern?

As a reminder, Heine *et al.* (2007) added several effect sizes to their meta-analytical Investigation 2. We explain, in Appendix II, how each of the added effect sizes is irrelevant to the question of whether self-other comparisons become more favourable to self with increasing importance of the comparison dimension. Nonetheless, we did code and include in the factorial design whether effect sizes from primary studies are relevant or not.

Method

Data sources. Data were obtained from: (i) the seven studies included in Investigation 2 of Sedikides *et al.*'s

Table 3 Characteristics of the 29 effect sizes included in Investigation 2

Wave	Study	Culture	Relevant	<i>r</i>	Standard error
1	Brown and Kobayashi (2002), Study 1	E	Y	0.2706	0.2236
1	Brown and Kobayashi (2002), Study 1	W	Y	0.2959	0.1768
1	Brown and Kobayashi (2002), Study 2	E	Y	0.1075	0.1715
1	Brown and Kobayashi (2002), Study 3	E	Y	0.1357	0.1155
1	Heine and Lehman (1997), Study 1	E	Y	0.4106	0.1155
1	Heine and Lehman (1997), Study 1	W	Y	0.3860	0.1179
1	Heine and Lehman (1999)	E	Y	0.0393	0.0814
1	Heine and Lehman (1999)	W	Y	0.0566	0.1111
1	Kobayashi and Brown (2003)	E	Y	0.2362	0.1400
1	Kobayashi and Brown (2003)	W	Y	0.2024	0.1336
1	Sedikides <i>et al.</i> (2003)	E	Y	0.3574	0.1491
1	Sedikides <i>et al.</i> (2003)	W	Y	0.3516	0.1491
2	Heine <i>et al.</i> (2001), Study 1	E	N	-0.2700	0.1170
2	Heine <i>et al.</i> (2001), Study 1	W	N	0.1900	0.1348
2	Heine <i>et al.</i> (2001), Study 2	E	N	-0.2100	0.1111
2	Heine <i>et al.</i> (2001), Study 2	W	N	0.2100	0.1280
2	Heine and Lehman (1999)	E	N	0.1000	0.0817
2	Heine and Lehman (1999)	W	N	-0.0390	0.1111
2	Heine and Renshaw (2002)	E	N	-0.3100	0.1459
2	Heine and Renshaw (2002)	W	N	0.0700	0.1348
2	Kitayama <i>et al.</i> (1997), Study 1	E	N	-0.2500	0.1291
2	Kitayama <i>et al.</i> (1997), Study 1	W	N	0.2100	0.1005
2	Kitayama <i>et al.</i> (1997), Study 2	E	N	-0.0800	0.0845
2	Kitayama <i>et al.</i> (1997), Study 2	W	N	0.1400	0.0909
3	Gaertner <i>et al.</i> (2006)	E	Y	0.3854	0.1324
3	Hamamura <i>et al.</i> (2006), Study 1	E	Y	0.2500	0.1890
3	Hamamura <i>et al.</i> (2006), Study 1	W	Y	0.3700	0.1644
3	Hamamura <i>et al.</i> (2006), Study 2	E	Y	0.3300	0.1098
3	Hamamura <i>et al.</i> (2006), Study 2	W	Y	0.4400	0.1562

Effect sizes are presented in the Pearson *r* metric and were transformed to Fisher *Z* prior to analysis. Standard errors are presented in the Fisher *Z* metric.

E, Easterners; N, no; W, Westerners; Y, yes.

(2005) meta-analysis, which yielded 12 effect sizes from 12 samples; (ii) the six additional studies included in Investigation 2 of Heine *et al.*'s (2007) meta-analysis, which yielded 12 effect sizes from 12 samples; and (iii) two studies by Hamamura, Heine, and Takemoto (2006), which yielded four effect sizes from four samples, as well as one study by Gaertner *et al.* (2006), which yielded one effect size from one sample. For expedience, we refer to these three sources of the 29 effect sizes as Wave 1, 2, and 3, respectively.

Participants from Western or Eastern samples ostensibly rated self and other on a series of attributes and rated the personal importance of each attribute. As displayed in Table 3, the 12 effect sizes introduced by Wave 2 of Heine *et al.* (2007) used alternative procedures and are irrelevant to the issue of whether self-other comparison becomes increasingly self-favouring as a function of personal importance of the comparison dimension (see also Appendix II).

Calculation of effect sizes. Authors of the primary studies graciously provided for each sample the average within-person correlation between the self-other rating and the importance rating. A positive correlation indicates that the self-other comparison became more favourable to self as the importance of the comparison dimension increased. A negative correlation indicates that the self-other comparison became more favourable to self as the importance comparison dimension decreased. A correlation of zero indicates the absence of a linear relationship between the self-other rating and the importance of the comparison dimension.

It is unclear, however, how to interpret correlations from the 12 irrelevant effect sizes. The components of those correlations vary and are consistent only in the fact that they do not involve a self-other comparison and subjective importance rating. The last two columns of Table 3 display the average within-person correlation and the correspond-

Table 4 Correlation between self-other comparison and attribute importance as a function of culture and relevance

Data Set	Relevant Yes				Relevant No				Culture × Relevant	
	Western	Eastern	<i>Q</i>	<i>p</i> =	Western	Eastern	<i>Q</i>	<i>p</i> =	<i>Q</i>	<i>p</i> =
Including Sedikides <i>et al.</i> (2003; Study 2)										
1	0.28*	0.29*	0.02	0.8875	0.17*	−0.12*	18.45	0.0000	9.32	0.0023
2	0.34*	0.23*	2.82	0.0931	0.13*	−0.19*	23.54	0.0000	4.48	0.0343
3	0.28*	0.23*	0.66	0.4166	0.17*	−0.19*	26.37	0.0000	10.29	0.0013
4	0.34*	0.29*	0.57	0.4503	0.13*	−0.12*	15.70	0.0001	3.82	0.0506
Excluding Sedikides <i>et al.</i> (2003; Study 2)										
1	0.27*	0.28*	0.03	0.8625	0.17*	−0.12*	18.45	0.0000	8.92	0.0028
2	0.34*	0.22*	2.81	0.0937	0.13*	−0.20*	23.54	0.0000	3.80	0.0513
3	0.27*	0.22*	0.59	0.4424	0.17*	−0.19*	26.37	0.0000	9.77	0.0018
4	0.34*	0.28*	0.60	0.4386	0.13*	−0.12*	15.70	0.0001	3.28	0.0701

* $p < 0.05$. *Q*-tests (d.f. = 1) of the Culture–Relevant interaction and the culture effect in levels of relevance were performed on Fisher Z transformed correlations. Tabled correlations are transformed from Fisher's Z and presented in the Pearson *r* metric.

ing standard error for each sample. We transformed the correlations with Fisher's Z prior to analysis.

Results and discussion

We created eight data sets to address two issues. As in Investigation 1, half of the data sets included the two effect sizes from Sedikides *et al.* (2003; Study 2), whereas the other half excluded them. We further varied those data sets to avoid violating the independence assumption. As Table 3 displays, the Eastern and Western effect sizes from Heine and Lehman (1999) in Wave 1 (effect sizes = 0.0393 and 0.0566) are derived from the same samples as the corresponding effect sizes in Wave 2 (effect sizes = 0.1000 and −0.039). We implemented two strategies to form data sets. The first strategy was to bracket the best approximation of the effect by excluding from Set 1 the two smallest non-independent effect sizes (i.e. 0.0393 and −0.039) and excluding from Set 2 the two largest non-independent effect sizes (i.e. 0.0566 and 0.1). The second strategy was to exclude non-independent effect sizes based on whether they were coded as being relevant to the issue. Set 3 excluded the two effect sizes coded as irrelevant (i.e. 0.1 and −0.039), and Set 4 excluded the two effect sizes coded as relevant (i.e. 0.0393 and 0.0566).

We regressed effect sizes on culture, relevant and Culture × Relevant. We planned to use a mixed-effects model. However, analysis of each data set evidenced a non-significant *Q*-test of the homogeneity assumption. Consequently, we report results of the fixed-effects model. As displayed in Table 4, each data set yielded the same conclusion.

In particular, analysis of each data set yielded a Culture–Relevant interaction. Studies yielding effect sizes relevant

to whether self-other comparison varies with the importance of the comparison dimension did not produce an effect of culture. Both cultures evidenced significantly positive correlations of equivalent magnitude indicating that Westerners and Easterners manifested self-other comparisons that increasingly favoured the self to the extent that the comparison dimension was personally important. In contrast, studies yielding effect sizes irrelevant to the issue produced an effect of culture. Westerners evidenced a significantly positive correlation, Easterners a significantly negative correlation. Unfortunately, those correlations are uninterpretable, given the inconsistency across effect sizes in the components of the correlation (i.e. each effect size addresses a different question).

Finally, readers might note the elevated *p*-values from analyses excluding Sedikides *et al.* (2003; Study 2). Importantly, those analyses yielded estimated effects equivalent to those that included Sedikides *et al.* The elevated *p*-values are due to the reduction in power incurred by the exclusion of additional effect sizes, rather than the values of those effect sizes.

In summary, the relevant studies boost the thesis of pancultural self-enhancement. Members of Western and Eastern culture self-enhance to the extent to which they subjectively deem as important the dimension of comparison. These findings replicate those of Sedikides *et al.* (2005; Investigation 2). As an aside, Heine *et al.* (2007) argued that the better-than-average effect 'artificially inflates estimates of the correlations between self-enhancement and importance for people of both cultures'. Even if this is the case, the point is inconsequential for our meta-analytical findings. The effect was significant for both cultural groups, and it was manifested selectively: For Westerners, the better-than-average judgement was corre-

lated with individualistic dimensions, for Easterners with collectivistic dimensions.

General discussion

Heine *et al.* (2007) suggested that the exclusion of six papers compromised the meta-analytical conclusion Sedikides *et al.* (2005) reached, namely that self-enhancement is pancultural. Furthermore, Heine *et al.* maintained that their meta-analysis, which included the six papers, demonstrated that self-enhancement is a Western motive because: (i) Easterners and Westerners did not differentially enhance on individualistic and collectivistic dimensions; and (ii) only Westerners enhanced to the extent to which the comparison attribute was important.

In response to Heine *et al.*'s (2007) challenge, we carried out a new meta-analysis, in which we included those six papers. We identified several faults in their meta-analytical procedure, as described in the Results and Discussion sections of our Investigations 1 and 2. More importantly, we noted a serious omission: their meta-analysis was misspecified. In particular, Heine *et al.* failed to take into account two moderators central to the theory. These are: (i) whether the comparison dimensions were validated or not; and (ii) whether effect sizes were relevant to the self-enhancement by personal importance association. Our theoretical formulation predicts that pancultural self-enhancement will be observed in the case of validated and personally important dimensions.

We addressed the Heine *et al.* (2007) procedural faults and omissions in two investigations. In Investigation 1, we demonstrated that Westerners self-enhance on individualistic dimensions, whereas Easterners self-enhance on collectivistic dimensions. This finding replicates that of Sedikides *et al.* (2005; Investigation 1). In Investigation 2, we demonstrated that both Westerners and Easterners self-enhance on personally important dimensions. This finding also replicates that of Sedikides *et al.* (2005; Investigation 2). In summary, the existing evidence is consistent with the view that tactical self-enhancement is pancultural.

End notes

1. We were able to include in our meta-analysis the Markus and Kitayama (1991) study thanks to Heine *et al.* (2007) providing an estimate of the effect size.
2. Although there are several conceptual and methodological reasons for excluding the Heine and Lehman (1995) study (see Appendix I), we nevertheless decided to include it in order to rule out further dispute of a 'selection bias' on our behalf.
3. In the Gaertner *et al.* (2006) study, 60 Taiwanese participants rated themselves as superior to their peers on (7) collectivistic

rather than (7) individualistic traits, while regarding the former traits as more personally important than the latter.

References

- Becker, B. J. (2006). The meaning and suitability of various effect sizes for structured rater X ratee designs. *Psychological Methods, 11*, 72–86.
- *Brown, J. D. & Kobayashi, C. (2002). Self-enhancement in Japan and America. *Asian Journal of Social Psychology, 5*, 145–167.
- Field, A. P. (2001). Meta-analysis of correlation coefficients: A Monte Carlo comparison of fixed- and random-effects methods. *Psychological Methods, 6*, 161–180.
- Field, A. P. (2003). The problems using fixed-effects models of meta-analysis on real-world data. *Understanding Statistics, 2*, 77–96.
- *Gaertner, L., Sedikides, C. & Chang, K. (2006). *Tactical Self-Enhancement in Taiwan: Implications for Psychological Health*. Unpublished manuscript. Knoxville, TN: University of Tennessee.
- *Hamamura, T., Heine, S. J. & Takemoto, T. (2006). *Why the Better-Than-Average Effect Is a Worse-Than-Average Measure of Self-Enhancement*. Unpublished manuscript. Vancouver, BC: University of British Columbia.
- Hedges, L. V. & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods, 3*, 486–504.
- Heine, S. J. (2005). Where is the evidence for pancultural self-enhancement?: A reply to Sedikides, Gaertner, and Toguchi (2003). *Journal of Personality and Social Psychology, 89*, 531–538.
- *Heine, S. J. & Lehman, D. R. (1995). Cultural variation in unrealistic optimism: Does the West feel more vulnerable than the East? *Journal of Personality and Social Psychology, 72*, 1268–1283.
- *Heine, S. J. & Lehman, D. R. (1997). Culture, dissonance, and self-affirmation. *Personality and Social Psychology Bulletin, 23*, 389–400.
- *Heine, S. J. & Lehman, D. R. (1999). Culture, self-discrepancies, and self-satisfaction. *Personality and Social Psychology Bulletin, 25*, 915–925.
- *Heine, S. J. & Renshaw, K. (2002). Interjudge agreement, self-enhancement, and liking: Cross-cultural divergences. *Personality and Social Psychology Bulletin, 28*, 578–587.
- Heine, S. J., Kitayama, S. & Hamamura, T. (2007). Inclusion of additional studies yields different conclusions: Comment on Sedikides, Gaertner & Vevea (2005). *Journal of Personality and Social Psychology, 10*, 49–58.
- *Heine, S. J., Kitayama, S., Lehman, D. R., *et al.* (2001). Divergent consequences of success and failure in Japan and North America: An investigation of self-improving motivations and malleable selves. *Journal of Personality and Social Psychology, 81*, 599–615.
- 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.

- Hornsey, M. J. & Jetten, J. (2005). Loyalty without conformity: Tailoring self-perception as a means of balancing belonging and differentiation. *Self and Identity*, 4, 81–95.
- Hunter, J. E. & Schmidt, F. L. (2000). Fixed effects vs random effects meta-analysis models: Implications for cumulative knowledge in psychology. *International Journal of Selection and Assessment*, 8, 275–292.
- Johnson, B. T. & Eagly, A. H. (2000). Quantitative synthesis in social psychological research. In: H. T. Reise & C. M. Judd, eds. *Handbook of Research Methods in Social and Personality Psychology*, pp. 496–528. New York: Cambridge University Press.
- *Kitayama, S., Markus, H. R., Matsumoto, H. & Norasakkunkit, V. (1997). Individual and collective processes in the construction of the self: Self-enhancement in the United States and self-criticism in Japan. *Journal of Personality and Social Psychology*, 72, 1245–1267.
- *Kobayashi, C. & Brown, J. D. (2003). Self-esteem and self-enhancement in Japan and America. *Journal of Cross-Cultural Psychology*, 34, 567–580.
- *Kurman, J. (2001). Self-enhancement: Is it restricted to individualistic cultures? *Personality and Social Psychology Bulletin*, 12, 1705–1716.
- *Markus, H. R. & Kitayama, S. (1991). Cultural variation in the self-concept. In: G. R. Goethals & J. Strauss, eds. *Multidisciplinary Perspectives on the Self*, pp. 18–48. New York: Springer-Verlag.
- Morris, S. B. & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, 7, 105–125.
- National Research Council (1992). *Combining Information: Statistical Issues and Opportunities for Research*. Washington, DC: National Academy Press.
- Neiss, M. B., Stevenson, J., Sedikides, C., Kumashiro, M., Finkel, E. J. & Rusbult, C. E. (2005). Executive self, self-esteem, and negative affectivity: Relations at the phenotypic and genotypic level. *Journal of Personality and Social Psychology*, 89, 593–606.
- *Norasakkunkit, V. & Kalick, S. M. (2002). Culture, ethnicity, and emotional distress measures: The role of self-construal and self-enhancement. *Journal of Cross-Cultural Psychology*, 33, 56–70.
- Pyszczynski, T., Greenberg, J., Solomon, S., Arndt, J. & Schimel, J. (2004). Why do people need self-esteem?: A theoretical and empirical review. *Psychological Bulletin*, 130, 435–468.
- *Ross, M. Heine, S. J. Wilson, A. E. & Sugimori, S. (2005). Cross-cultural discrepancies in self-appraisals. *Personality and Social Psychology Bulletin*, 31, 1175–1188.
- *Sedikides, C. Gaertner, L. & Toguchi, Y. (2003). Pancultural self-enhancement. *Journal of Personality and Social Psychology*, 84, 60–70.
- Sedikides, C. Gaertner, L. & Vevea, J. L. (2005). Pancultural self-enhancement reloaded: A meta-analytic reply to Heine (2005). *Journal of Personality and Social Psychology*, 89, 539–551.
- Singelis, T. M. (1994). The measurement of independent and interdependent self-construals. *Personality and Social Psychology Bulletin*, 20, 580–591.
- Taylor, S. E. Lerner, J. S. Sherman, D. K. Sage, R. M. & McDowell, N. K. (2003). Portrait of the self-enhancer: Well-adjusted and well-liked or maladjusted and friendless? *Journal of Personality and Social Psychology*, 84, 165–176.
- Vevea, J. L. & Hedges, L. V. (1995). A general linear model for estimating effect size in the presence of publication bias. *Psychometrika*, 60, 419–435.
- Vevea, J. L. & Woods, C. M. (2005). Publication bias in research synthesis: Sensitivity analysis using a priori weight functions. *Psychological Methods*, 10, 428–443.
- Yamaguchi, S. Greenwald, A. G. Banaji, M. R., et al. (in press). Apparent universality of positive implicit self-esteem. *Psychological Science*.

References accompanied by an asterisk include studies that were used in the meta-analysis.

Appendix I

Exclusion rationale for Investigations 1 and 2 of Sedikides et al. (2005)

I. Inclusion criteria for Investigation 1 of Sedikides et al. (2005). Studies had to: (i) identify cultural membership on the basis of nationality, ethnicity, or self-construal; (ii) provide a measure of self versus other perception; and (iii) assess the self-other comparison on individualistic versus collectivistic dimensions.

II. Inclusion criteria for Investigation 2 of Sedikides et al. (2005). Criteria (i) and (ii) were the same as in Investigation 1. Criterion (iii) required studies to report the personal importance of the comparison dimension.

III. The six papers in question.

1 Heine and Lehman (1995). Assessed event expectations instead of self-defining attributes. Study 1 data did not involve event importance ratings. Also, the individualistic/collectivistic events of Study 2 were problematic, because they systematically differed on whether the consequences of the event befell the self or a close other: ‘individualistic versus collectivistic’ events were confounded with the target of the event’s outcome. For example, the ostensibly individualistic event ‘Sometime in the future you will get AIDS’ is suffered by the participant, whereas the ostensibly collectivistic event ‘Sometime in the future a member of your family will be convicted of an embarrassing crime’ is suffered by a member of the participant’s family. In contrast, all of our Investigation 1 studies held constant the target (i.e. self) of the individualistic/collectivistic attributes. Finally, the comparison dimension was not validated.

- 2 Heine and Lehman (1999). We excluded these data in Investigation 1, because no distinction was made between individualistic and collectivistic attributes. However, we did include these data in Investigation 2 (table 9, p. 546).
- 3 Heine *et al.* (2001). None of the four studies met inclusion criteria for either Investigation 1 or 2: none of the studies assessed perception of self versus other, while the studies did not distinguish between individualistic and collectivistic dimensions.
- 4 Heine and Renshaw (2002). The study met neither Investigation 1 inclusion criteria (no distinction was made between individualistic and collectivistic attributes) nor Investigation 2 inclusion criteria (the importance ratings were provided by others rather than the participant). Also, measures of 'accuracy/agreement' (comparing a participant's rating of self with the average rating that participant received by group members) rather than 'enhancement' were used. This practice differs from our question of whether a given person perceives the self more or less favourably than others, a question that necessitates comparing the rating of self with the rating of other.
- 5 Kitayama *et al.* (1997). The self-other ratings were obtained from two different sets of participants (i.e. self data from Study 1, 'other' data from Study 2) and did not reflect a given person's perception of self versus other. Also, no distinction was made between individualistic and collectivistic dimension, and dimension importance was not assessed ideographically.
- 6 Markus and Kitayama (1991). This study was not available from psycINFO.

Appendix II

Why the effect sizes that Heine et al. (2007) added are irrelevant to the 'importance' question

1. Heine and Lehman (1999). Sedikides *et al.* (2005) formed an effect size using participants' ratings of self versus other. Heine *et al.* (2007) included another effect size based on the standard deviation of an actual/ideal self rating task. Heine *et al.* formed a correlation between actual/ideal discrepancy and importance. This measure,

however, does not compare self with other and is, thus, irrelevant.

2. Heine *et al.* (2001; Study 1). This study does not involve self-other perception. Participants received success versus failure feedback and subsequently repeated the task. The authors measured persistence time on the second task. The reported correlation reflects the relation between persistence and perceived task importance.

3. Heine and Renshaw (2002). There were two problems with this study. First, the measure of self-enhancement is actually a measure of self-other accuracy. Participants interacted in five-person groups and completed trait ratings of self and the four other members. The effect size reflects the difference between one's self-rating and the average rating the self received by the four group members (i.e. self-rating-others' rating of self). Thus, the difference score does not address whether participants perceived themselves more positively than they perceived others. Rather, this score addressed whether participants perceived themselves differently than they were perceived by others. Second, the importance ratings with which the self-other score was correlated were obtained from persons other than the participants and, consequently, the ratings do not reflect subjective importance. (A pilot group of participants previously rated trait importance.)

4. Kitayama *et al.* (1997; Studies 1 and 2). Participants rated the self-esteem relevance of each of 400 situations (obtained by asking other samples of Japanese and Americans to list success vs failure situations). Participants in Study 1 rated the extent to which their own self-esteem would increase or decrease. Participants in Study 2 rated the extent to which the self-esteem of a typical undergraduate student would increase or decrease. In neither study did participants additionally rate the importance of the situations. The effect size from Study 1 reflects the extent to which one's own self-esteem rises and falls in response to Japanese versus American situations. The effect size from Study 2 reflects the extent to which the self-esteem of a typical undergraduate student rises and falls in response to Japanese versus American situations. Neither effect size compares self versus other. Also, neither effect size involves a correlation with subjective importance.