

The Translation Equivalence of Two Japanese Versions of the Rosenberg Self-Esteem Scale

By R. A. Brown

Abstract

The present research investigated the comparability of two commonly used Japanese translations of the Rosenberg Self-esteem Scale (RSES). Study 1 revealed that participants evaluated themselves more negatively on one version than the other. Study 2 however showed that participants did not evaluate a target person who was depicted endorsing the RSES items significantly differently regardless which version was used.

Key words: self-esteem, Japan, Rosenberg Self-Esteem Scale, translation equivalence

Many studies of Japanese self-esteem (SE) have been conducted and most have used the Rosenberg (1965) Self-Esteem Scale (see Heine & Hamamura, 2007, and Heine, Lehman, Markus, & Kitayama, 1999 for reviews). Several Japanese translations of the Rosenberg Self-Esteem Scale (RSES) are available (Hori, 2003; Mimura & Griffiths, 2007). Inevitably, different translations will convey different albeit possibly subtle shades of meaning, which may influence participants' responses, a point recently noted by Schmitt & Allik (2005). Few studies provide details about the translation used, but the two most widely used would appear to be those of Hoshino (1970) and Yamamoto, Matsui, and Yamanari (1982). Recent research (Brown, in press) using the Hoshino version raised the possibility that the results were more negative than they might have been had the Yamamoto et al. version been used. The purpose of the present research was to ascertain whether the Hoshino translation in fact elicits more negative global self-appraisals.

Study 1

Method

Participants. Participants were 168 Japanese university students (89 males, 78 females, one unspecified, average age = 19.2, $SD = 1.3$), from two universities in the Tokyo area. Participants were first year college students enrolled in English classes designed to meet graduation requirements and the minimum standards of proficiency set by the national government (Hashimoto, 2007), accordingly enrollment in these classes does not indicate any special familiarity with, or interest in, Anglo-American culture or psychological theories. One group was significantly older, $t(164) = 8.88, p < .0001$ but exploratory analyses revealed that the two groups did not differ with regard to any of the variables of interest, hence were aggregated for subsequent analyses. In both cases, questionnaires were filled out in large classes, voluntarily, and without compensation.

No deception or manipulation was involved.

Instrument. Both Hoshino and Yamamoto et al., translations of the RSES were used and were distributed randomly to participants. Slightly more than half (89 versus 79) filled out the Yamamoto et al version, but the difference was not significantly different from 0.5.

Effect size was estimated with Cohen's *d* (Cohen, 1988), which is the standardized mean difference between the two measurements. Following Cohen's (1992) suggestions, effect sizes of 0.2 or less are interpreted as small, 0.5 as medium, and 0.8 or above as large. Because probability levels are sensitive to sample size, lack of significance does not necessarily entail lack of an effect (Cohen,1990; Gelman and Stern, 2006), therefore effect sizes were calculated even when the $p < .01$ level was not reached.

Results

Age was positively correlated with one RSES item on the Yamamoto et al. version ("I feel that I have a number of good qualities"), $r(88) = .35, p < .001$, and with the summated RSES, $r(88) = .28, p < .01$, but did not correlate with any item on the Hoshino version. Both versions had Cronbach's alphas of .80 and both versions would have had alphas of .82 had the same item ("I wish I could have more respect for myself") been deleted. The aggregate RSES score was significantly more negative relative to the scale midpoint when the Hoshino version, $t(78) = 2.63, p \leq .01$, was used compared to the Yamamoto et al., version, $t(89) = 0.99, ns$.

Effect sizes were generally small and only in the case of one item ("all in all, I'm satisfied with myself") was it moderate.

Table 1. Means and Standard Deviations, Cohen's *d* and effect size *r*.

Item	Hoshino		Yamamoto et al		<i>d</i>	<i>r</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
1.	2.76	(0.96)	3.15	(0.86) *	.43	.21
2.	2.38	(0.95)	2.00	(0.88) *	.42	.20
3.	2.96	(0.94)	3.39	(0.92) *	.46	.22
4.	2.81	(1.09)	2.96	(1.02)	.14	.07
5.	3.15	(1.10)	3.12	(0.89)	.03	.02
6.	2.59	(1.19)	3.20	(1.02) **	.55	.26
7.	3.29	(1.04)	3.47	(1.12)	.17	.08
8.	3.29	(0.89)	3.43	(0.84)	.15	.07
9.	2.70	(1.19)	2.84	(1.27)	.11	.06
10.	2.20	(1.09)	3.06	(0.92) **	.85	.39
RSES	2.81	(0.63)	3.06	(0.59) *	.41	.20

Note. ** Row means differ significant at $p \leq .001$, * significant at $p < .01$

Note. Item 1 = I feel that I have a number of good qualities; item 2 = I wish I could have more respect for myself (reverse scored), item 3 = I feel that I am a person of worth, at least on an equal plane with others, item 4 = I feel I do not have much to be proud of (reverse scored), item 5 = I take a positive attitude toward myself, item 6 = I certainly feel useless at times (reverse scored), item 7 = All in all, I'm inclined to feel that I am a failure (reverse scored), item 8 = I am able to do things as well as most other people, item 9 = At times I think I am no good at all (reverse scored), item 10 = On the whole, I am satisfied with myself.

Discussion

Interpretation of the results are straightforward. The Hoshino version elicits more negative responses than the Yamamoto version does. Thus, while use of the Hoshino version should not vitiate conclusions that Japanese means are generally lower than American means on the RSES, it may overstate the negativity of Japanese SE.

Study 2

Method

Participants. Participants were 101 Japanese college students (65 males, 36 females) with an average age of 18.9 ($SD = 0.71$). Data from an additional 120 students from a previous study described below were used for comparison, bringing the total sample size to 221.

Instrument. Participants read a statement purportedly made by a hypothetical high SE (HSE) and low SE (LSE) target persons (TPs) of unspecified gender and age. The statements consisted of the five positive and five negative RSES items, respectively, slightly edited to form a coherent paragraph. The Yamamoto et al translation was used. In a recent previous study, involving 120 demographically similar students, Brown (in press) used the Hoshino translation. Data from that study were therefore available for comparison. Participants were asked first to assess the individual on 13 characteristics deemed relevant to the study (adapted from Brown, in press), namely, self-confidence, self-respect, self-esteem, likeability, capability, psychological health, adjustment, probability of future successful life outcomes (hereafter, success), need for psychological counseling (hereafter, need), narcissism, arrogance, humility, and depression. The 13 traits were aggregated into four clusters, based on correlational and semantic considerations (described in Brown, in press), labeled SELF-REGARD (self-esteem, self-respect, and self-confidence), ADJUSTMENT (psychologically healthy, needs counseling (reverse scored), depressed (reverse scored), and well-adjusted), HUMILITY (humility, arrogant [reverse scored] and narcissistic [reverse scored]), and SUCCESS (successful, capable, and likeable). Because pre-testing has demonstrated that few 19 year old Japanese university students are familiar with the English expression *self-esteem*, a description of self-esteem was provided (adapted from R. A. Brown, 2008-a). Assessments were made on appropriately labeled seven-point scales. The questionnaire was composed entirely in Japanese by the author with the assistance of a professional native speaker translator.

Results

Participants appraised the HSE TP higher in ADJUSTMENT, $t(212) = 2.63, p = .009$, and the LSE TP as lower in SELF-REGARD, $t(219) = 3.52, p = .001$ in the Yamamoto et al version compared to the Hoshino version. Effect sizes however were small. The other six clusters did not significantly differ. Two approached significance (HSE SUC and LSE ADJ), but even in these cases effect sizes were small. Means and standard deviations are shown in Table 2.

Table 2. Means, standard deviations, Cohen's *d* and effect size *r*.

	Hoshino Version		HSE Target Yamamoto et al. Version		d	r
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
HSE Target						
SR	5.34	(1.17)	5.63	(0.95)	.27	.13
ADJ	4.47	(0.97)	4.91	(0.90) *	.47	.23
HUM	2.87	(1.19)	3.19	(0.97)	.29	.14
SUC	3.77	(1.04)	4.11	(0.98)	.37	.16
LSE Target						
SR	2.76	(1.26)	2.21	(0.99) **	.48	.23
ADJ	3.43	(0.97)	3.13	(0.94)	.31	.15
HUM	5.25	(1.29)	5.52	(1.23)	.21	.11
SUC	3.29	(1.09)	3.12	(0.97)	.16	.08

* Row means differ at $p < .01$, ** $p < .001$.

In terms of the specific traits, when the Yamamoto et al. translation was used, the HSE target was viewed as significantly ($p < .01$) less arrogant, less in need of counseling, and more capable. However, the LSE target was viewed less positively in the Yamamoto et al., version, specifically as lower in confidence, self-esteem, and psychological health.

Discussion

Participants' general impressions of the HSE TP were similar irrespective of which translation was used. There were two rather small differences with respect to specific trait clusters, but with countervailing impacts on participants' impressions of the TP. Overall then, the two versions may be approximately equivalent when used to assess third parties (rather than the self).

General Discussion

As is well known, the exact wording of questionnaire items can influence participants' responses (Converse & Presser, 1988; Schwarz, 1999; Schwarz & Oyserman, 2001.) Variant translations of a single source instrument necessarily differ in phrasing and as linguists say, any difference in expression presupposes a difference in meaning (Brown, 1987; Bynon, 1977). In the present research, participants responded more self-critically to one of the two RSES translations under examination. But they evaluated a third person target approximately equally over both versions. Thus for some, but not all research purposes, the two versions can be regarded as equivalent to each other.

Much of cross-cultural psychological research involves the translation of self-report instruments. Psychologists are not usually experts in linguistics or translation and perhaps some important issues have tended to be glossed over. Studies generally report that a back-translation procedure was used to ensure accuracy, but seldom are the details of the process discussed. When unrelated languages are involved, arriving at semantically and functionally equivalent translations can be problematic, as it cannot always be safely assumed that indigenous concepts have exact counterparts in cultures that have historically evolved

independently, (Wierzbicka, 2004), even when a good deal of superficial lexical borrowing has occurred (Brown, 2008-b). Brown 2007 discusses some of these issues, such as pronoun choice. The Japanese language (among others) has a complex system of personal pronouns, each of which encodes different sorts of information about the relationships between interlocutors and context, none of which is precisely synonymous with the English pronouns (Hasegawa & Hirose, 2005; Martin, 1988; Maynard, 2007; Ono & Thompson, 2003). Similarly, concepts that have one set of implications in one context may have different implications in another context. One example is that of "being satisfied with oneself." For some Japanese participants, this expression suggests laziness, narcissism, and arrogance rather than healthy self-esteem (Brown, 2007; Brown, 2008-c; Brown, in press). Syntactic patterns may also contain ambiguities in the target language not present in the original. "I wish that I could..." is one such example, also discussed by Brown (2007), and which has been identified in a number of studies involving such languages as Japanese, Korean, and Chinese as having poor inter-item correlations with the remaining nine RSES items (Ang, Neubronner, & Leong, 2006; Cheng & Hamid, 1995; Farruggia, Chen, Greenberger, Dmitrieva, & Macek, 2004; Feather, & McKee, 1993; Hamid & Cheng, 1995).

Sireci, Yang, Harter, and Ehrlich (2006) recommend statistical evaluation to verify that a translation retains the functional meanings of the original. The same can be applied to competing translations of a source instrument. Detection of differential item functioning may indicate that one item may be closer in meaning to the original. It would not however say anything about which one. In the present case, it may be that Hoshino is too negative, but may equally be that Yamamoto et al., is too positive. Careful examination of both by a competent bilingual who understands the intent of the original is required. In the present case it becomes evident that Hoshino probably mistranslated the item "on the whole, I'm satisfied with myself." The phrase "on the whole" in English means "generally" or "everything considered" and Yamato et al., translated it accurately as "*daitai ni oite*." Hoshino however translated it as "*subete no ten de*," which actually means "in all ways." It is not surprising that participants would more often tend to reject the latter version. Apart from this, Hoshino is more "formal." Eight of the RSES items in the Hoshino translation begin with the phrase "*watashi/watakushi wa*," (customarily translated into English as "as for me") which undoubtedly focuses greater attention on the self. In Japanese, first-person pronouns are customarily omitted in colloquial speech and informal writing (Hasegawa & Hirose, 2005). Incorporating them calls special attention to the speaker by explicitly excluding other people as subjects of the predication. In contrast, the Yamamoto et al., translation eschews this phrase, giving it the tone or register of spoken rather than written Japanese. It is not unlikely that participants adopt a self-referential stance primed by the language of the research instrument. As Maynard (2007) points out, in Japanese there are a variety of first person forms of self-reference, each appropriate to a different situation, each expressing a different "self" and Japanese individuals have a variety of selves to choose from and enact. Nevertheless, the fact that alternative versions of the RSES differ from each other does not tell us which is closer to the original.

Recommendation

Given that Hoshino contains a mistranslation, it would probably be wise to use the Yamamoto et al., version, while not ignoring the possibility that it may present an excessively positive picture of participants' SE.

References

- Ang, R. P., Neubronner, S-A. O., & Leong, V. (2006). Dimensionality of Rosenberg's Self-Esteem Scale among normal-technical stream students in Singapore. *Current Psychology: Developmental · Learning · Personality · Social*, 25, (2), 120-131.
- Brown, R. A. (1987). Kyooyo and multigraphia in Japan. *Asian & Pacific Quarterly*, 19, 19-30.
- Brown, R. A. (2007). Remarks on Feigned Modesty, Language, and Japanese Self-Esteem *Information & Communication Studies*, 36, 25-34. Bunkyo University: Chigasaki, Japan.
- Brown, R. A. (2008-a). American and Japanese Beliefs about Self-Esteem. *Asian Journal of Social Psychology*, 4, 293-299..
- Brown, R. A. (2008-b). The Equivalence of Self-Esteem and *Jisonshin Asian Journal of Social Psychology*, 4, 300-304.
- Brown, R. A. (2008-c). Censure avoidance and self-esteem in Japan. *Journal of Social Psychology*, 148, 653-666.
- Brown, R. A. (in press). Perceptions of psychological adjustment, achievement outcomes, and self-esteem in Japan and America. *Journal of Cross-Cultural Psychology*.
- Bynon, T. (1977). *Historical linguistics*. Cambridge: Cambridge University Press.
- Cheng, S-T., & Hamid, P. N. (1995). An error in the use of translated scales: The Rosenberg Self Esteem scale for Chinese. *Perceptual & Motor Skills*, 81, 431-434.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.
- Cohen, J. (1990). Things I have learned (so far). *American Psychologist*, 45, 1304-1312.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155-159.
- Converse, J. M., & Presser, S. (1986). *Survey Questions: Handcrafting the Standardized Questionnaire*. Newbury Park, CA: Sage.
- Farruggia, S.B., Chen, C., Greenberger, E., Dmitrieva, J., & Macek, P. (2004). Adolescent self-esteem in cross-cultural perspective: Testing measurement equivalence and a mediating model. *Journal of Cross-Cultural Psychology*, 35, 719-733.
- Feather, N.T., & McKee, I. R. (1993). Global Self-Esteem and Attitudes Toward the High Achiever for Australian and Japanese Students *Social Psychology Quarterly*, 56, 65-76.
- Gelman, A., & Stern, H. (2006). The difference between "significant" and not significant" is itself not significant. *The American Statistician*, 60, 328-333.
- Hamid, H. P., & Cheng, S-T. (1995). To drop or not to drop an ambiguous item: A reply to Shek. *Perceptual & Motor Skills*, 81, 988-990.
- Hasegawa, Y., & Hirose, Y. (2005). What the Japanese language tells us about the alleged Japanese relational self. *Australian Journal of Linguistics*, 25, 219-251.
- Hashimoto, K. (2007). Japan's language policy and the "lost decade. In A B. M. Tsui & J W. Tollefson (Eds.). *Language policy, culture, and identity in Asian contexts* (pp. 25-36). Mahwah, NJ: Lawrence Erlbaum.
- Heine, S. J., & Hamamura, T. (2007). In search of East Asian self-enhancement. *Personality and Social*

Psychology Review, 11, 4-27.

- 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.
- Hori, K. (堀 啓造). (2003). Rosenberg 日本語訳自尊心尺度の検討 Retrieved November 12, 2005 from <http://www.ec.kagawa-u.ac.jp/~hori/yomimono/sesteem.html>
- Hoshino, A. (1970). Affect and education II. *Child Psychology*, 24, 161-193. [Japanese].
- Martin, S. E. (1988). *A reference grammar of Japanese*. Rutland, VT: Charles E. Tuttle.
- Maynard, S. K. (2007). *Linguistic creativity in Japanese discourse*. Amsterdam: John Benjamins.
- Mimura, C., & Griffiths, P. (2007). A Japanese version of the Rosenberg Self-Esteem Scale: Translation and equivalence assessment [Abstract]. *Journal of Psychosomatic Research*, 62, 589-594. Retrieved November 13, 2008, doi:10.1016/j.jpsychores.
- Ono, T., & Thompson, S. A. (2003). Japanese (w)atashi/ore/boku 'I': They're not just pronouns. *Cognitive Linguistics*, 14, 321-347.
- 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.
- Schwarz, N. (1999). Self-Reports: How the questions shape the answers. *American Psychologist*, 54, 93-105.
- Schwarz, N., & Oyserman, D. (2001). Asking questions about behavior: Cognition, communication, and questionnaire construction. *American Journal of Evaluation*, 2, 127-160.
- Sireci, S. G., Yang, Y., Harter, J., & Ehrlich, E. J. (2006). Evaluating guidelines for test adaptations: A methodological analysis of translation quality. *Journal of Cross-Cultural Psychology*, 37, 557-567.
- Wierzbicka, A. (2004). 'Happiness' in cross-linguistic & cross-linguistic perspective. *Daedalus*, (spring), 34-43.
- Yamamoto, M., Matsui, Y., & Yamanari, Y. (1982). *Ninchi sareta jiko no shosokumen no kouzou*. [structure of self-recognition]. *Japanese Journal of Educational Psychology*, 30, 64-68.

Author Note

I am indebted to Yoko Kondo and Harada Shinichi for assistance with various aspects of the research. Correspondence concerning this article should be addressed to R.A. Brown, 1-2-22 East Heights # 103, Higashi Kaigan Kita, Chigasaki-shi 253-0053 Japan. Electronic mail may be sent via internet to RABrown_05@hotmail.com.