

Scale translation from source to target language as the creation of psychometrically equivalent parallel forms

LEONARD I. JACOBSON, J. ANTONIO HERNANDEZ, and JOSE L. GARCIA
University of Miami, Coral Gables, Florida

An investigation was conducted in order to test the hypothesis that scale translation may be conceptualized productively as a process in which psychometrically equivalent or parallel forms of a source scale are generated in desired target languages. The 190 subjects were administered the Anderson-Jacobson Beliefs about Equal Rights for Men and Women Scale, in either English or Spanish. The latter scale was constructed according to exacting standards, so as to create a parallel form of the English version. All of the subjects were native speakers of the language in which they participated. It was found that women demonstrated greater belief in equal rights than men did, in both languages. Analyses of differences in language of administration, point biserial correlations between items and total score, and patterns of endorsement among equivalent items were made for both languages. These analyses fully confirmed the original hypothesis.

An enduring controversy in psycholinguistics concerns the most desirable approach to be employed in translating scales from one language to another. Sophisticated techniques have been employed to overcome the problems in moving from the source to the target language (Brislin, 1970, 1976; Hulin, 1987).

An alternative approach is to consider the problem in traditional psychometric terms, as the creation of functionally equivalent forms of a scale in more than one language, so that source and target scales share common psychometric properties. When viewed in this manner, the problem becomes one of creating parallel forms of a single test, in which both forms are equivalent with regard to fundamental psychometric properties. As Gulliksen (1950) has shown, these basic properties include the equivalence of means, standard deviations, and correlations among the forms:

$$X_1 = X_2 = X_3 = X_4 = \dots = X_K \quad (1)$$

$$S_1 = S_2 = S_3 = S_4 = \dots = S_K \quad (2)$$

$$r_{12} = r_{13} = r_{14} = r_{23} = \dots = r_{KK} \quad (3)$$

In the present study, a paradigmatic effort at achieving these objectives was implemented. An investigation was performed in which a parallel form of a previously validated English-language scale was generated in Spanish. A series of objectives was established for the relationship of the Spanish-language scale to the English-language one. The objectives were as follows:

1. The total scores should not differ significantly for both tests.

2. Means and standard deviations between the two scales should not be significantly different.

3. The correlation of each item with total score should be similar for both forms.

4. As a corollary of the third objective, the percentage of subject endorsement for each item should be similar in both languages.

5. Subject variables, such as sex, should result in comparable scores on both forms, with regard to both total score and interitem patterns of analysis. That is, significant differences between men and women may be found, but they should be equal in both languages.

METHOD

Subjects

There were 190 subjects, including 76 men and 114 women. All of the subjects were volunteer students at the University of Miami. There were 137 subjects in Condition 1 and 53 in Condition 2.

Experimental Conditions

In order to participate, a subject had to meet one of two sets of conditions. In Condition 1, the subjects were required to have spoken English as their first language. In order to participate in Condition 2, the subjects had to have been born in a Spanish-speaking country and to have Spanish as their first language. The stringent requirements of Condition 2 were necessary, because at the University of Miami there are many students from Spanish-speaking homes who were born in the United States but are woefully inadequate in reading and writing Spanish. It was deemed desirable to have the Spanish speakers at the same level of fluency in that language as the English speakers were in their language.

The English-Language Scale

The scale adapted into Spanish was the Anderson-Jacobson Beliefs about Equal Rights for Men and Women scale (BAERS; Jacobson, Anderson, Berletich, & Berdahl, 1976). This well-validated scale, which consists of 28 items, measures the extent to which individuals hold congruent attitudes with regard to the rights and gender roles of men and women. The BAERS is different from other sex-role measurement devices, in that it does not measure attitudes toward men and/or women

Requests for reprints should be sent to Leonard I. Jacobson, Psychology Department, University of Miami, P.O. Box 248185, Coral Gables, FL 33124.

directly and quantitatively, but rather measures the congruence of attitudes toward both sexes.

The test requires the examinees to compare their attitudes toward the prerogatives of both sexes when evaluated simultaneously. Because it is generally believed that North Americans and Latins differ greatly in their perceptions of sex roles, this scale was considered an excellent test case with respect to whether or not it is possible to create personality scales in different languages whose cultures' norms are greatly different as well.

Generation of the Spanish Form

Three judges well-educated in the Spanish language were each requested to translate the BAERS into Spanish, following these rules: (1) The Spanish employed was to be at about the 10th-grade level—as is the case with the language of the English version; and (2) regionalisms and expressions unfamiliar to any major Spanish-speaking country were to be avoided.

The committee then met and compared their versions in order to arrive at a single unified form that was considered the best version possible according to unanimous vote. This scale was then administered to 72 subjects, and revisions were made on the basis of the results obtained, in order to achieve greater psychometric equivalence among the items.

Procedure

The subjects filled out a demographic data sheet and then the BAERS. The subjects in Condition 1 received the scale in English; the subjects in Condition 2 received the scale in Spanish.

RESULTS AND DISCUSSION

An analysis of variance was performed to determine the effects of language of administration and sex on BAERS scores. No significant effect of language of administration was found [$F(1,186) = 0.01, p > .05$], nor was there any interaction of language of administration \times sex [$F(1,186) = 0.14, p > .05$]. However, a significant effect of sex was in evidence [$F(1,186) = 15.50, p < .0001$].

On the English version, males scored 18.55 ($SD = 4.90$) and females scored 21.49 ($SD = 4.52$). On the Spanish version, males scored 18.78 ($SD = 5.48$) and females scored 21.11 ($SD = 4.67$). It is evident that the

scores obtained for the English and Spanish forms are similar, and that men believe less strongly than women in equal rights, regardless of language. In addition, point biserial correlations between the items and total scores yielded similar patterns of intercorrelation for both the English and Spanish forms, as well as similar patterns of intercorrelation for the two tests for men and for women. Further chi-square analyses of the percentage of subjects endorsing each item on the two forms indicated no significant differences between the English and Spanish forms for any item (all $ps > .05$).

In summary, a Spanish-language equivalent form of the BAERS was created with comparable total scores, standard deviations, item endorsements, and point biserial correlations between items and total score. Therefore, we may conclude that the five formulated objectives were attained, and that psychometrically equivalent or parallel forms were indeed created. These results confirm the validity of the present approach, and they indicate that scale translation may be conceptualized successfully as a process in which psychometrically equivalent forms of a source scale are generated in desired target languages.

REFERENCES

- BRISLIN, R. (1970). Back-translation for cross cultural research. *Journal of Cross-Cultural Psychology*, **1**, 185-216.
- BRISLIN, R. (1976). Translation research and its applications: An introduction. In R. Brislin (Ed.), *Translation: Applications and research*. New York: Wiley/Halstead.
- GULLIKSEN, H. (1950). *Theory of mental states*. New York: Wiley.
- HULIN, C. L. (1987). A psychometric theory of evaluations of item and scale translations. *Journal of Cross-Cultural Psychology*, **18**, 115-142.
- JACOBSON, L. I., ANDERSON, C. L., BERLETICH, M. S., & BERDAHL, K. W. (1976). Construction and initial validation of a scale measuring beliefs about equal rights for men and women. *Educational & Psychological Measurement*, **36**, 913-918.

(Manuscript received January 30, 1989.)