Using Translated Instruments In Research

Judith E. Beauford, University of the Incarnate Word, USA
Yosuke Nagashima, University of the Incarnate Word, USA
Ming-Hsun Wu, University of the Incarnate Word, USA

ABSTRACT

International and cross cultural research in the global community often requires translations of instruments. This paper reviews the development of a translation protocol for instruments written in English for a Western culture and used for a different language and culture.

Keywords: translation, instrument, cross-cultural research, cultural sensitivity

NEED INTERNATIONAL RESEARCH

The “need for global exchange” called for by Schmieding and Kokuyama (1995, p. 820) and Maneesriwongul and Dixon (2004) for research is apparent in all areas of study, both because of the global interchange of commerce and the improvements that can be accomplished worldwide as we learn from one another. The globalization of higher education includes cross-cultural research projects often conducted by teams from many countries and the presence of students from many countries in American universities. As these students prepare theses and dissertations they naturally and appropriately want to apply their research efforts to topics of interest in their home countries. They often choose research instruments and techniques designed for Western cultures to be applied to their native culture.

Many of the instruments designed for use in research in the social sciences have established records in English speaking cultures (Mason, 2005). Using them rather than developing a similar instrument for the target population often saves time, money and effort. Hambleton and Kanjee (1993) note that the technical expertise for development of a new instrument may not be available in the target culture. Translation also allows for cross-culture comparative studies. In some instances, the use of translated instruments facilitates testing in one’s primary language in an attempt for equity among students from several language heritages (Hambleton & Kanjee).

CULTURAL APPROPRIATENESS

Every culture has unique values, organizational systems, and environments. Awareness of these is essential to understanding national, ethnic, racial, linguistic groups. Cultural sensitivity, deeper understanding, and respects other cultures are necessary for valid translation and cross-cultural research. Differences in value structures and social mores can dramatically affect how a sample responds to a survey, questionnaire, or other research instrument even with the best translations possible. Questions about personal opinion, social habit, and even business practice that might seem innocuous to one culture can be seen as highly offensive to another. For example, a primary difference observed by international students is the American sense of individualism which can be seen in sharp contrast to the group-orientation of many cultures. Confucian respect for authority results in uneasiness in employees when asked to evaluate their leaders. Hofstede (2001) explains “In some cultures and subcultures, being polite to the other person is more important than supplying objectively correct information; in some, respondents will never use no” (p. 21). It would be considered as rude. This might impact how Japanese participants, for example, would respond to a questionnaire asking for an opinion, perception, or preference.

“Language and translation of research instruments are crucial, as is the matching of samples from the different cultures for functional equivalence” (Hofstede, 2001, p. 1). Hierarchical structure in business and educational institutions can radically change to roles and responsibilities of persons in what seems to be similar position. For example, the principal in an American school may play a much different role from the head of a cram
school in Taiwan. High power-distance in some cultures influence the appropriateness of some leadership behaviors that seem quite natural in cultures with low power-distance.

**TRANSLATION PROTOCOL IN THE LITERATURE**

When the need for translation is known at the time of construction of the instrument, care can be taken to facilitate conversion to the second language. “Choice of item formats, stimulus material for the instrument, vocabulary, sentence structure, and other aspects which might be difficult to translate well can all be taken into account in the initial instrument development to minimize problems later in translation” (Hambleton & Kanjee, 1993, p. 6). Taking both languages into consideration in the writing process, words that have no equivalent are avoided (Hambleton & Kanjee, 1993; Chang, 1994). Care must be taken to assure clear purpose and content meaning (Yu, Lee & Woo, 2004; Chang, 1994), perhaps by adding context to difficult terms (Chang, 1994; Hofstede, 2001). Simple sentences in the active voice are most easily translated.

Translation strives to achieve conceptual equivalence (Schmieding & Kokuyama, 1995; Mason, 2005). “Conceptual equivalence implies that an item may be translated into different words, but the original meaning or conceptual framework remains intact” (Mason, 2005, p. 70). The translation strives for “equivalence of item meanings and idioms in both cultures” (Yu, Lee & Woo, 2004, p. 310). But “the wider apart the structures of the two languages, the less simple the task” (Hofstede, p. 21).

Efforts to clarify meaning can introduce unwanted results. Hofstede (2001) and others recommend the “transfer [of] desired meaning through circumlocution” (Hofstede, p. 21). But “[v]aried cultural experiences and worldviews” (Mason, 2005, p. 69) confound the problematic use of idiom and context. Yu, Lee, and Woo (2004) list common errors to include “(a) distorting the meaning of a word or phrase, (b) adding extra meaning to the original intent, and (c) erroneously deleting the meaning of some words in the [source language] version” (p. 314). Hambleton and Kanjee (1993) cite test format as a source of difficulty with the translation of instruments. “[I]t cannot be assumed that all students are as familiar with multiple choice items as U.S. students” (p. 4). The issues must be addressed in both cross-cultural research and intercultural research using translated instruments, to avoid differences found being attributed to groups when they are actually caused by errors in translation (Schmieding & Kokuyama, 1995).

The literature is fairly consistent in its recommendations for translation strategies. Many suggest that multiple translators be used in the process (Hambleton & Kanjee, 1993; Maneesriwongul & Dixon, 2004) and that translators be most familiar with the target language and culture (Hambleton & Kanjee; Schmieding & Kokuyama, 1995; Hofstede, 2001). They should be familiar with the subject matter (Hambleton & Kanjee; Chang, 1994) and trained in testing and scale construction (Hambleton & Kanjee). While it would be rare for all translators available to a project to possess all these characteristics, a mixture where each strength is represented works well. Using a committee approach, the translation is developed and then tested for equivalence (Hambleton & Kanjee, 1993; Maneesriwongul & Dixon, 2004).

Hambleton and Kanjee (1993) recommend checking for equivalence of frequency counts of words. For instruments with strong test-retest reliability, a pilot test-retest with bilingual participants using the two versions (Hambleton & Kanjee; Maneesriwongul & Dixon, 2004; Yu, Lee & Woo, 2004) can be useful if time is allowed between administration of the two versions to assure that the responses to the second are in response to the questions rather than recall of the first version (Griffee, 2001). Many use interview or focus groups of bilingual persons (Chang; Griffee, 2001; Hofstede, 2001; Maneesriwongul and Dixon, 2004) to review each item and reach consensus on the best translation.

“Multiple techniques should be used in all cross-cultural research. . . . [When] “semantic equivalence cannot be achieved, researchers should consider replication of the instrument development process in the target language, rather than using an instrument which will not make sense in the target language” (Maneesriwongul & Dixon, p. 183).
ESTABLISHING VALIDITY AND RELIABILITY

The literature is not as consistent in recommendation for the establishment of validity of translated research instruments by statistical methods. With a compilation of recommendations from several sources and personal experience a strategy evolves.

“Validation is the offering of evidence that the persons answering the questionnaire items understood what the items were asking in a way reasonably the same as the questionnaire developer” (Griffee, 2001, p. 2) and consequently a measure of the extent to which an instrument measures what it intends to measure. Reliability is a measure of the stability of result for equivalent subjects. “Reliability and validity of the source language version should not necessarily be assumed for the target language version” (Maneesriwongul & Dixon, 2004, p. 182). When any change is made to a research instrument, whether survey, questionnaire, attitudinal measure, cognitive measure or other, the result is a new instrument that must be tested for validity and reliability (Griffee). Mason (2005) suggests that even without change such instruments must be validated for each new study population.

A pilot study with participants from the population should be conducted to establish descriptive statistics and Cronbach alphas for each item and the whole which are then compared with the original version of the instrument (Mason, 2005). Application of the instrument in both the original and translated versions to bilingual participants should result in strong correlations as would be expected in other test-retest studies.

Factor analysis is indicated when “[t]here is a possibility that an instrument with adequate internal consistency may have a different underlying structure than the original version” (Mason, 2005, p. 71). Principle components analysis is appropriate when “exploratory in nature,” confirmatory factor analysis “when particular underlying factors are assumed to be indicated by previous studies” (p.71). Mason recommends a sample size of 20 participants per variable with “caution …for small sample sizes such as 100” (p. 71). Again results should be compared with the original version of the instrument.

OUR EXPERIENCES

In use of translated instruments for both cross-cultural and intercultural research, we have experienced most of the drawbacks and have experimented with most of the translation strategies mentioned in the literature. Assuming that if asked the same question, bilingual participants would give the same answers in both languages, Pai (2006) used a small (n=10) pilot study to check for translation accuracy. In his instruments measuring health-related self-efficacy, physical activity identity, and stress, he found good correlation between English and Chinese versions of the instruments. J. Wu (2006) did not, however, have the same experience. Her instrument was a measure of comfort level with a particular course and she found that the answers varied widely with whether the survey was given in the first or second language of her pilot sample. She concluded as Chang (1994) and Hofstede (2001) had that comfort with the language being used influences comfort levels reported for the course. Seeing this result, M. Wu (2006) devised a method of using his bilingual pilot sample as recommended by Chang (1994), Griffee (2001), Hofstede (2001) and Maneesriwongul and Dixon (2004) to come to consensus about the translations of his instruments measuring perceived leadership styles and effectiveness in higher education. Ko (2006) used a well known leadership style instrument published in both English and Chinese on a sample of 753 students in Taiwan and found a difference between the versions in the underlying constructs using a principle components analysis. With appropriate response to these translation validation efforts followed by statistical measures of reliability and validity, interpretations of results from these studies were reported with care.

UIW TRANSLATION PROTOCOL

Our practices continue to develop. At this point the recommended protocol is as follows.

1. An initial translation is evaluated by a committee of bilingual professionals in the field of study of the research. Requirements for familiarity with the culture and field are usually met. Expertise in scale construction is rare. The translation is amended according to their suggestions until there is consensus about the accuracy and validity of the instrument.
2. The instrument is tested on a small focus group of 3 to 10 bilingual participants as close to the study population as possible. Item analysis and interviews with this focus group result in further amendment of the instrument as is warranted.
3. A larger pilot study is run with a sample of 20 to 50 from the target population and descriptive and reliability statistics are compared with the published results of the original instrument. Only egregious problems with these results produce change at this point.
4. The instrument is applied to the full sample in the research study.
5. The results are evaluated for reliability by the use of Cronbach’s alpha.
6. A principal components analysis is performed with each subscale of the instrument, checking for satisfactory loadings on each component within the subscale.
7. Items, subscales, and factors that do not reach satisfactory levels for the Cronbach’s alpha and principal components analysis are removed from further analysis.
8. Cronbach alphas, principal components analysis, and occasionally structural equation modeling are used to establish validity and reliability of the resultant instrument for analysis of research questions.

AUTHOR INFORMATION

Yosuke Nagashima holds Master Degree in International Education and is completing his dissertation for a Ph.D. in the Dreeben School of Education at the University of the Incarnate Word in San Antonio, TX, with a concentration in International Education and Entrepreneurship. He is from Tokyo, Japan, has lived in several countries through study abroad programs since age 14, and has been in the U.S. for 11 years. He has worked as a medical and cultural interpreter for the City of San Antonio and as an instructor of U.S. Study Abroad at Japan College of Foreign Language in Tokyo.

Ming Hsun Wu holds a Ph.D. from the Dreeben School of Education at the University of the Incarnate Word with a concentration in Organizational Leadership. In his dissertation he used the Competing Values Framework in his investigation of leadership behaviors of presidents in technology universities. He is a faculty member in the International Business School at Chien Kuo Technology University, Taiwan.

Judith E. Beauford holds a Ph.D. from the University of Texas at Austin. She is a Professor of Mathematics and Director of Doctoral Studies in the Dreeben School of Education at the University of the Incarnate Word, San Antonio, TX. Dr. Beauford has published work in Mathematics Education, particularly on the integration of mathematics and science and the influence of language on mathematics learning.

REFERENCES


