

Test Translation Matters

No. 2, February 2007

Test Translation Matters (TTM) is an occasional publication of Second Language Testing, Inc. (SLTI). The purpose of TTM is to inform educators, testing specialists, and the public on matters and issues concerning the translation or adaptation of standardized tests. TTM is archived on the SLTI website, www.2LTI.com, so that readers can search back numbers for information on issues of problems they are encountering. We realize that our readers' time is limited, so each number is short, dealing only with one

Back translation

The number of school-age students in US schools who speak a language other than English at home grew rapidly, from 3.8 million in 1979 to 9.9 million in 2004, according to the National Center for Education Statistics. There are also more than 3 million children currently classified as limited English proficient. The No Child Left Behind Act (NCLB) mandates the "inclusion of limited English proficient students, who shall be assessed in a valid and reliable manner and provided reasonable accommodations on assessments ... including, to the extent practicable, assessments in the language and form most likely to yield accurate data on what such students know and can do in academic content areas." Therefore, translation of statewide content area assessments is on the rise.

One concern states face is ensuring the accuracy of the translations, and verifying that the translated versions of the assessments measure the same knowledge, skills and abilities (KSAs) as the original (English language) assessments.

This is of particular concern when the target languages involved are less commonly taught languages, as often states' Department of Education personnel are simply unable to read the resulting products. Nevertheless, it is always a concern whenever a test is translated, even when Spanish, the most common language for test translation, is involved.

One potential verification tool, back translation, has been the subject of debate in the literature on translation. Back translation essentially entails re-translating a document from the target language back into the source language, on the premise that if the original translation was accurate, the translation back into the source language should be nearly identical to the original document.

Cross-cultural psychologist Richard Brislin was the first to write extensively about back translation as a way to ensure a quality translation of a test instrument (1970; 1976; 1986; Brislin, Lonner, & Thorndike, 1973). As a

or two specific matters in test translation. The focus, at least initially, is test translation and adaptation in the context of K-12 state assessment programs. This issue of TTM was prepared by Charles Stansfield of SLTI and Melissa Bowles of the University of Illinois. Comments on the topic are welcome. If you do not wish to receive further issues of TTM, simply send an email to Cstansfield@2LTI.com. Anyone wishing to be added to the list should send a request to the same address.

cross-cultural psychologist, Brislin was interested in examining the consistency of traits and constructs across cultures. As described in the literature on cross-cultural research, back translation involves asking a bilingual to translate the original test to the target language and then having a different bilingual translate it back to English. The two English versions are then compared, and points of disagreement are used to identify problems in the initial translation. The forward translation is then modified accordingly. The process of creating a back translation and comparing it with the source document is repeated until the source document and back translation agree (Marin & Marin, 1991).

Back translation was also discussed in an article by Ron Hambleton (1993, p. 62) prepared for the Third International Math and Science Study (TIMMS) and in *The Guidelines for Adapting Educational and Psychological Tests* (Hambleton, 1994). These documents were widely circulated in the psychometric community where some interpreted the discussion as an endorsement of back translation. In fact, although Hambleton indicates that back translation can be helpful, although he also points out many flaws associated with using the procedure. Indeed, the TIMMS did not use back translation because of the flaws and the expense involved (Maxwell, 1996). In subsequent presentations at national conferences, Hambleton has clearly stated that he does not endorse back translation. Nonetheless, back translation continues to be viewed by some in the psychometric community as the appropriate methodology for creating a test translation or ensuring the equivalence of the translated test.

While back translation is viewed in the test translation literature as a method for drafting, reviewing, and revising a translation, its primary function is to identify and correct errors in the forward translation. However, relying on the back translation to examine the quality of the translated document is problematic. First, the lack of agreement

between the original document and the back translation may be due to problems with the back translation, not to problems with the initial forward translation. That is, the back translation is as likely to contain translation errors (mistranslations, omissions, insertions) as is the forward translation. Hence, once the back translation has been conducted, the test program administrator now has two translations but no verification of the quality of either.

Second, a translator's knowledge that the initial forward translation will be validated by a back translation may influence his or her approach to forward translation. By producing a very literal forward translation, the translator can ensure that the back translation will produce a document that is very similar to the source document. Unfortunately, such a translation is likely to produce stilted rather than natural expression and result in a test that is difficult to read and hence less accessible to the examinee population.

Normally, other methods, such as successive reviews and revisions of a forward translation by other professional translators and by representatives of the examinee population, are available to identify and correct problems with the forward translation. For this reason, professional translation agencies do not normally use back translation as a quality control procedure unless specifically required to do so by their clients. Instead, they rely on successive review and revision of the forward translation by reviewers of known competence as translators.

Nevertheless, some states have used back translation in the development of state assessments for ELLs. Minnesota, for instance, experimented with back translation to obtain feedback on its translation of one of several state assessments to Spanish (Liu et al., 1999). A full year was devoted to the translation and pilot testing of the test. In the real world of state assessment, however, only two to four weeks typically are available for translation, review, revision, typesetting, proof, and correction of a test. Although back translation produced a good Spanish translation in Minnesota, Liu et al. noted the high cost and length of the process. Successive iterations of forward translation and revision would work just as well, and the translation equivalence of the two versions can be assured more quickly and at less cost. In subsequent years, Minnesota did not use back translation to verify its translated assessments, but rather forward translation with iterative rounds of review and revision.

Stansfield and Kahl (1998) evaluated several matters (including forward translation) during the pilot testing of the Spanish version of the Massachusetts (MCAS) Mathematics assessments in 1997. They found that the *iterative approach* to forward translation produced a translation that was well received by examinees and teachers representing different nationalities. The iterative approach involved a 16-step process, which included successive reviews and revisions of the forward translation

by professional translators. Stansfield and Auchter (2001) reported on the successful translation of new forms of the GED test battery, which involved iterative revisions of a forward translation, followed by multiple verifications of the test key for the translated versions.

Because back translation is frequently mentioned in the academic literature it is sometimes requested by purchasers of translation services. Because of this, professional translators sometimes debate the ethics of doing a back translation, implying that an ethical translator will first try to explain to the client that a back translation may not be the most direct way to ensure a good translation, and it is more costly. One translation company (Barinas, 2007) includes it on a page devoted translation myths and misinformation. An officer of another company (Crystal, 2007) has published an extensive discussion of the flaws of back translation on a website for translators, in an effort to help translators educate their customers.

Based on thirteen years of experience in test translation, Second Language Testing, Inc. recommends and regularly employs the technique of iterative review and revision, rather than back translation, to ensure the equivalence of translated assessments. State education agencies and other parties considering back translation as a quality control measure are encouraged to consider the pros and cons discussed in this issue of Test Translation Matters before making any decisions.

References

- Barinas. (1970). Translation myths and misinformation. Retrieved on February 4, 2007 from <http://www.barinass.com/myths.htm>
- Brislin, R. (1970). Back translation for cross-cultural research. *Journal of Cross Cultural Psychology, 1*, 185-216.
- Brislin, R. (1976). *Translation: Applications and research*. New York: Gardner
- Brislin, R. (1986). The wording and translation of research instruments. In W.J. Lonner & J.W. Berry (Eds.), *Field methods in cross-cultural psychology*. Newbury Park, CA: Sage.
- Brislin, R., Lonner, W.J. & Thorndike, R.L. (1973). *Cross-cultural research methods*. New York: John Wiley.
- Cristal, S.M. (2007). Same questions, different continent. Retrieved on February 4, 2007 from <http://www.translationdirectory.com/article1150.htm>
- Hambleton, R.K. (1993). Translating achievement tests for use in cross-national studies. *European Journal for Psychological Assessment, 9*, 1, 57-68.
- Hambleton, R.K. (1994). Guidelines for adapting educational and psychological tests: A progress report. *European Journal for Psychological Assessment, 10*, 3, 229-244.
- Liu, K.K., Anderson, M.E., & Swierzbis, B. (1999). *Bilingual accommodations for limited English proficient students on statewide reading tests: Phase 1*. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.
- Marin, G. & Marin, B.V. (1991). *Research with Hispanic populations*. Newbury Park, CA: Sage.
- Maxwell, B. (1996). Translation and cultural adaptation of survey instruments. In M.O. Martin & D.L. Kelly (eds.), *Third International Math and Science Study technical report, volume 1: design and development*. Chestnut Hill, MA: Boston College.
- Stansfield, C.W. & Kahl, S.R. (1998). *Lessons learned from a tryout of Spanish and English versions of a state assessment*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.
- Stansfield, C.W. & Auchter, J.E. (2001). A process for translating achievement tests. In C. Elder & A. Brown (eds.), *Experimenting with uncertainty*. Studies in language testing #11, 73-80. Cambridge: Cambridge University Press.