

TRANSLATION OF RESEARCH INSTRUMENTS: RESEARCH PROCESSES, PITFALLS AND CHALLENGES

K.A. Dhamani, PhD candidate
Faculty of Nursing
University of Alberta, Canada

M.S. Richter, D Cur
Faculty of Nursing
University of Alberta, Canada

Corresponding author: solina.richter@ualberta.ca

ABSTRACT

Multilingual and multi-ethnic societies are becoming the norm in the era of globalisation. Given the cultural diversity and multiple languages spoken in many countries, healthcare researchers (including nurses) are challenged to use psychometrically sound research instruments that are culturally and linguistically sensitive. Most psychometrically sound research instruments have been developed and their properties assessed in English-speaking populations. A literature review was performed to understand the process of translation, use of qualitative and quantitative methods to assess the quality of translation, and lastly, to identify strategies to overcome the challenges of the translation process. One-way translation was observed to be the most utilised method. Translation methods and processes have many challenges, but applying relevant strategies could reduce errors and pitfalls.

KEYWORDS: challenges, cultural sensitivity, equivalence, instrument translation, semantic similarity, translation pitfalls, translation processes

INTRODUCTION AND BACKGROUND INFORMATION

Multilingual and multi-ethnic societies are becoming the norm in the era of globalisation (Capitulo, Cornelio & Lenz, 2001). Cultural beliefs and practices of immigrants can influence health and health seeking behaviours. Healthcare providers are challenged not only to assess and provide appropriate healthcare to these diverse populations, but also to include them in health services research (Carlson, 2000; Duffy, 2006; Hilton & Skrutkowski, 2002). Nurse researchers should become aware of the details of translation methodology (Carlson, 2000) and report translation procedures, including statistical analyses, used in the development of cross-cultural research instruments. Most psychometrically sound research instruments have been developed and their

properties assessed in English-speaking populations (Capitulo et al., 2001; Duffy, 2006; Hall, Wilson & Frankenfield, 2003). Given the cultural diversity and multiple languages spoken in many countries, healthcare researchers are challenged to either develop or translate existing research instruments to be culturally and linguistically sensitive (Capitulo et al., 2001; Hilton & Skrutkowski, 2002). Accurate cross-cultural instruments assist in discriminating the true variance because of true differences in the phenomena being studied. Interventions, according to Wang, Lee and Fetzer (2006:310) which are ...“effective in one culture can be tested and applied in a different culture”.

Translating research instruments for cross-cultural and multi-ethnic relevance and from one language to another is a complex process (Duffy, 2006). According to Hilton and Skrutkowski (2002), becoming familiar with the basic problems of language equivalence, cultural constructs, and psychometric changes built into the translation process is important in developing a culturally equivalent translated instrument. Health researchers including nurses should therefore understand the complexity involved in the translation process and minimise the methodological pitfalls. The scope of this article is not about presenting a systematic review or a comprehensive review but to demonstrate authors' understanding of the instrument translation process. As such, no particular criteria or search strategy was used to select the articles for the review; but authors were cognisant in choosing a relevant range of articles on the topic. The purpose of this article is to describe different types of translation methods and the process of scientific translation for a research instrument. One of the main aims is also to discuss the pre-testing methods utilised for linguistic validation and equivalence of translated instruments; and to discuss some of the common challenges with translation and highlight methods to minimise errors in translation.

TYPES OF TRANSLATION METHODS AND THE TRANSLATION PROCESSES

According to the National Child Traumatic Stress Network ([NCTSN] 2006), the purpose of translation is to accurately capture the meaning of the source instrument by using appropriate language that the target group can easily understand. Similar to the findings by Maneesriwongul and Dixon (2004) of the literature review of 47 studies on instrument translation, no standard guidelines were found which are used by researchers for instrument translation. The literature reviewed, described mainly three types of translation methods: one-way or expert translation, committee or focus group approach, and double or back translation. Researchers usually combine these methods to develop quality translation of research instruments, but at times they use only one method (Capitulo et al., 2001; Duffy, 2006; Eremenco, Cella, & Arnold, 2005; Hilton & Skrutkowski, 2002). A rigorous version of the double or back translation method is preferred to other methods in ensuring valid and reliable data (Eremenco et al., 2005).

One-way or expert translation

The one-way or expert translation method is most frequently used because of its low cost and simplicity (Carlson, 2000; Hilton & Skrutkowski 2002). With this method, a bilingual person with demonstrated knowledge and expertise in the target language, and familiarity with the content area, accomplishes the instrument translation from the original language to the target language (Capitulo et al., 2001). The disadvantages of this method include the limited availability of experts and the difficulties of establishing the validity of experts. Complete reliance on the sole translator's knowledge and skills may lower validity and reliability of an instrument (Capitulo et al., 2001; NCTSN, 2006).

Committee or focus group approach

The method of translation by a committee or a focus group committee entails a group of bilingual individuals who accomplish the translation of an instrument from the source to a target language by working independently or collaboratively as a group. The group meets to reconcile discrepancies if necessary (Carlson, 2000; Martinez, Marin & Schoua-Glusberg, 2006; NCTSN, 2006). This method does not include back-translation (Martinez et al., 2006). When an instrument with a target language is to be developed for a multinational clinical trial, native speakers from a variety of countries or areas are included, using the same committee approach. Developing a single translation that is appropriate for use in different regions is termed a universal translation approach (Capitulo et al., 2001; Eremenco et al., 2005). The feasibility of the universal translation approach is subject to criteria including that subject matter should be devoid of cultural bias, common symptoms and the nature of the instrument in question. (Eremenco et al., 2005). The translated version of the instrument by committee approach captures the conceptual meaning of the item and the appropriateness of the language (NCTSN, 2006). Consensus among the bilingual group may yield a more accurate version than the subjective opinion of a single translator in the one-way translation method. However, the process is time consuming, and building consensus may be difficult because of the varying degrees of expertise in the subject matter and translation (Carlson, 2000; Hilton & Skrutkowski, 2002).

Double or back translation

The double or back-translation method is the most reliable method of attaining semantic equivalence between the source language and the target language (Capitulo et al., 2001; Duffy, 2006; Eremenco et al., 2005; Hilton & Skrutkowski, 2002; Wang et al., 2006). As described in figure 1, the first step involves one or more bilingual experts translating the instrument from the source language into the target language (forward translation). Secondly, one or more bilingual experts independently translate the forward-translated instrument back into the source language (back translation or double translation) without knowledge of the original instrument (Duffy, 2006; Eremenco et al., 2005; Hilton & Skrutkowski, 2002; NCTSN, 2006; Ramirez, Teresi, Holmes, Gurland, & Lantigua, 2006). Lastly, either the researcher consults with both translators

or forward and back translators, compare and modify their original and back-translated versions. Discrepancies are addressed to produce a final version of the translation that is conceptually and linguistically equivalent (Carlson, 2000; Hilton & Skrutkowski, 2002). The focus of forward translation is on retaining the conceptual meaning of the item rather than on producing simply a literal, word-for-word translation, whereas capturing the literal meaning of the item is the focus of back or double translation (Eremenco et al., 2005; Ngai, Chan, & Holroyd, 2007). Carlson mentioned that forward translation is more of an instrument adaptation, whereas instrument translation as ‘translation’ and ‘adaptation’ differ in meaning based on literal and conceptual forms of a new instrument (2000).

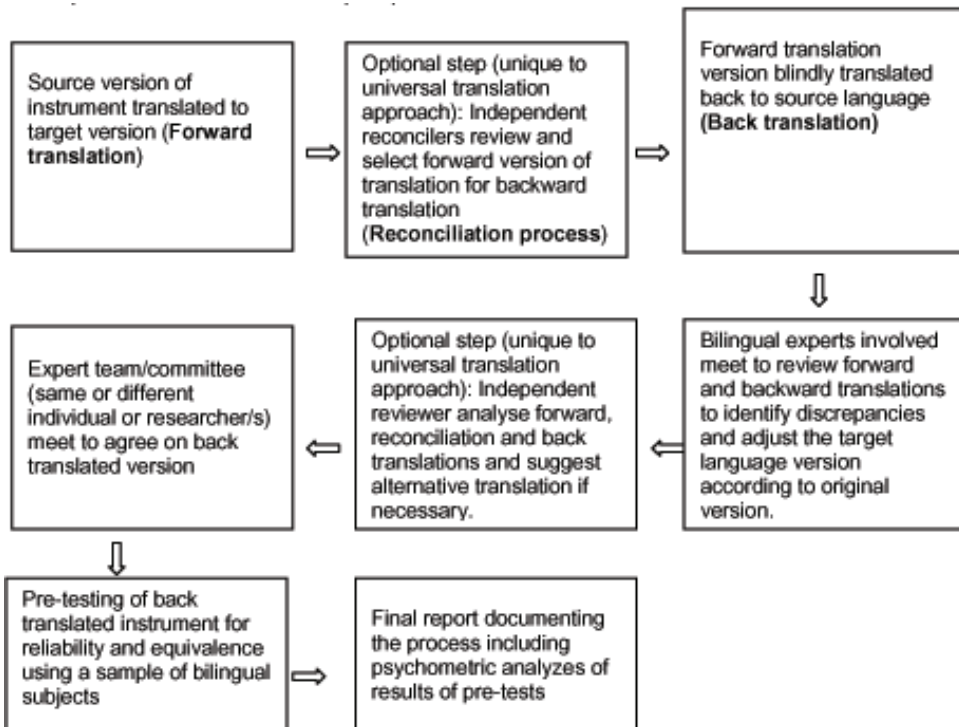


Figure 1: Steps in forward and backward translation

Two optional steps may be added to the above method: reconciliation and an independent review process. In the reconciliation process, one or more reconcilers review forward versions of the translation and select the most appropriate version for back translation. They can also alter the forward translation to make it more suitable if required. This step enhances the possibility of linguistic compatibility with different dialects and introduces objectivity (Eremenco et al., 2005; Ramirez et al., 2006). In the independent review process, two or more bilingual experts independently analyse the forward translations, the reconciliation version, and the back translation and comment or suggest alternative translations, if necessary. Eremenco et al. (2005:222) claimed that multiple reviews

are vital for a universal translation approach and a step that makes this methodology superior to others. They also maintain that the purpose of independent reviews is “to analyze the target-language translation in light of all the information provided and select the best target-language translation for each item”.

Eremenco et al. (2005) also used a language coordinator (an experienced translator) and independent proofreaders in their translation methodology, whereas in the other articles that we reviewed, the researchers did not identify their use as part of the translation process. According to Eremenco et al. (2005), the role of the language coordinator is to produce a final translation of the item and proofread the final version for grammatical and formatting errors. An independent proofreader double-proofreads the final version to detect errors before the validation process is initiated. The translation and back translation methods involve all or a combination of forward translation, a reconciled version of the forward translation, blind back translation, and either consensus from a committee or independent reviews by experts (Eremenco et al., 2005; Ramirez et al., 2006). Although preferred and widely used, this method is time consuming, expensive, and impractical for multilingual studies. Differences between the original and back-translated version may still exist despite the translators’ care and diligence (Carlson, 2000; Duffy, 2006; Hilton & Skrutkowski, 2002).

Decentering process

Capitulo et al. (2001) identified *decentering* of the translation process as an extension of back translation. This process involves the steps of back translation. Once the final version has been developed after back translation, the individual or organisation that holds the copyright for the original instrument and the translation team/individual mutually agree on semantic equivalence between the original and translation instrument by modifying the original instrument with universally better understood language. Hilton and Skrutkowski (2002) describe symmetrical and asymmetrical categories of translation and noted that symmetrical translation requires that the original and translated instruments have loyalty of meaning and colloquialness and be equally familiar. Hence, decentering refers to the translation process in which both the source and the target language versions are considered equally important and subject to change during translation (NCTSN, 2006; Poss, 1999). In the asymmetrical translation category, the original language is retained without modification; thereby the translated version may present a literal translation of words and may lack conceptual equivalence (Eremenco et al., 2005). One of the reasons that an instrument developer or organisation resists making changes is to avoid compromising on the previously established validity and reliability of the instrument (Eremenco et al., 2005; Hilton & Skrutkowski, 2002).

The translation process is considered incomplete until the instrument is pretested with the members of the target culture. Pre-testing of an instrument serves mainly two purposes: “checks for translation quality and the practical aspects of test administration” (Hilton & Skrutkowski, 2002:3). It also helps to correct possible translation problems

that threaten the semantic equivalence of the final translation (Eremenco et al., 2005). Respondents for pre-testing are recruited based on their native-language abilities (Capitulo et al., 2001; Duffy 2006; Eremenco et al., 2005; Hilton & Skrutkowski, 2002). Qualitative methods of testing will establish translation that is equivalent to the source instrument; whereas quantitative methods will determine the measurement properties of the translated instrument. Therefore, applying both methods could be useful.

METHODS FOR TESTING TRANSLATED RESEARCH INSTRUMENTS

Qualitative methods

Qualitative methods yield a translation that is equivalent to the source instrument. Both retrospective debriefing interviews and cognitive debriefing interviews are the qualitative approaches utilised to measure the quality of the translations.

Retrospective debriefing interviews

The respondents receive the translated version of the questionnaire and are asked to complete it. The interviewer thereafter reviews the questionnaire for completeness and checks for missing data or other problems in general. In a retrospective debriefing interview, the interviewer asks the respondents if they found any items difficult to understand, irrelevant, or offensive and solicits their views on relevant topics that might need to be included in the questionnaire. Items that receive negative comments from a significant number of respondents are reworded based on the respondents' inputs. Proofreading is required after changes have been made to the questionnaire (Eremenco et al., 2005).

Cognitive debriefing interviews

This is an alternative approach used to identify the respondents' interpretation of the translations. The only difference between retrospective and cognitive debriefing interviews is that, in this approach, the respondents are asked questions on an item-by-item basis to paraphrase their understanding of each item and interpret items that they found problematic in translation. This approach ensures that the meaning intended in the source instrument is retained in the target instrument. However, the disadvantages are that it requires a great deal of probing to obtain the information and is time consuming because it involves indepth interviewing (Eremenco et al., 2005; Ramirez, 2006).

Quantitative methods

Quantitative methods are used to determine the measurement property of the translated instrument and because of this, it is critical that some level of quantitative analysis be performed. This also provides an "evidence of measurement equivalence and similar performance with the source instrument" (Eremenco et al., 2005:225). Internal

consistency analysis and differential item functioning are the quantitative approaches utilised to measure the quality of translations.

Internal consistency analysis

An instrument is said to have internal consistency reliability when all of its subparts measure the same attribute under study. This approach assesses measurement error in multi-item questionnaires (Polit, Beck & Hungler, 2001). Also, through internal consistency a new standard for interpreting responses is developed for a target language group, thus reducing cultural biases built into the instrument (Hall et al., 2003). If sufficient information has been gathered through pre-testing, the questionnaire as a whole and its subparts are analysed for comparison with the source version. Estimates of the reliability of translated instruments are usually computed by using Cronbach's alpha and obtaining descriptive statistics (mean, SD, range). An alpha level of 0.7 or above for a subscale and for a whole questionnaire is considered to indicate an acceptable reliability coefficient (Duffy, 2006; Eremenco et al., 2005; Polit et al., 2001; Willgerodt, Kataoka-Yahiro, Kim, & Ceria, 2005).

Validity

The measure of validity is another important criterion for evaluating a quantitative instrument. In the context of translation, measurement of construct validity is highly recommended. Construct validity determines the ability of the translation instrument to measure the concept that it is supposed to measure (Polit et al., 2001). The statistical procedure most widely utilised for measuring the construct validity of a translated research instrument is factor analysis, identifying clusters of related items on a scale of the instrument. Item subscale correlations below 0.2 reflect poorly performing items. These may result from errors in the translation. It then suggests the need to review the item translation (Eremenco et al., 2005).

Differential item functioning (DIF)

DIF compares an item's level between two different groups using the same instrument. According to Ramirez et al. (2006), groups who share a similar ethnic background and the same language may still show significant distinctive and cultural differences. DIF identifies items that perform differently for individuals or groups of different levels, different ethnicities, or different languages. "Differences in item content and administration may also affect comparisons across studies and interpretations of findings" (Ramirez, 2005:S98). Because DIF also identifies items that may cause measurement bias, excluding the item from the instrument or from future analysis may reduce threats to validity. DIF analysis also ensures "translation equivalence and justifies data pooling and reporting" (Eremenco et al., 2005: 226).

DEVELOPMENT OF A FINAL REPORT

Once the research team has completed the required psychometric tests and the analysis, the detailed report should be produced and include the translation methods and process; the testing process, the statistical analysis and changes to the translation based on pretesting (Duffy, 2006; Eremenco et al., 2005; Wang et al., 2006). Such detailed reports serve as a reference point for future translations and meet the demands of regulatory authorities.

CHALLENGES AND METHODS TO MINIMISE ERRORS IN TRANSLATION

One of the greatest challenges of translation is to retain the culturally appropriate and meaningful original items in the translated instrument (Wang et al., 2006). Hence, translation methodology should aim at producing a translation of the original instrument that is equivalent to the source version and vice versa (Eremenco et al., 2005). In this context semantic equivalence is the unbiased measurement between two translated instruments or the degree of similarity between the source and translated instrument (Eremenco et al., 2005; Polit et al., 2001). Combining different methods of translation and utilising qualitative and quantitative methodologies to test them enhances the linguistic and cultural equivalence of a translated version. Table 1 illustrates different types of equivalence as one of the ways of assuring validity of the source and the target language versions of the instrument (Peters & Passchier, 2006).

Table 1: Types and descriptions of equivalence

Type of equivalence	Description of equivalence
Content/Item	Content of each item is relevant in each culture. Back translation method, pre-testing of the instrument, cognitive debriefing interviews may be used to enhance equivalence (translating some constructs may still remain a challenge)
Semantic	Maintaining similarity in meaning of each item in each culture after translation. Native bilingual translators' performance of forward translation, the introduction of a reconciliation process, feedback from a committee, and field testing of the original and translated instruments with monolingual and bilingual participants may enhance the conceptual equivalence
Technical/Operational	Data-collection methods for the source instrument and the target version are comparable. Differences in administration or the formatting of an instrument may result in variations in response
Criterion	Interpretation remains the same as the norm for each culture. Retrospective or cognitive debriefing interviews, pre-testing of the instrument may enhance criterion equivalence

Conceptual	Instrument measures the validity of the concept in each culture. Internal consistency analysis and differential item functioning enhances conceptual equivalence
------------	--

(Bowden & Fox-Rushby 2003; Duffy 2006; Hilton & Skrutowski 2002; Peters & Passchier, 2006)

The equivalence of a translation instrument is achieved by addressing the methodological pitfalls inherent in translating questionnaires for cross-cultural research (Hilton & Skrutkowski, 2002). Instrument translation is a major methodological issue, and nurse and healthcare researchers need to adequately describe translation methods and processes (Willgerodt et al., 2005). Instrument developers and teams face many challenges in developing translations, despite choosing the right method, utilising multistage procedures, having bilingual and bicultural experts working independently or in committees, employing language coordinators, and utilising qualitative and quantitative methods to test translations remain difficult to attain 100% equivalence (Duffy, 2006; Eremenco et al., 2005; Hilton & Skrutkowski, 2002).

In spite of involving bilingual and bicultural experts in the translation process, there is still a possibility that limited linguistic and cultural views will be represented, which in turn may influence the quality of the translation. If experts do not have a pan-cultural view of the target language, there is a risk of limited generalisability. Furthermore, the bilingual experts' varying degrees of expertise in the subject matter and translation may also result in translations that are incongruent and inconsistent compared to the original instrument. The committee members and/or bilingual experts face another challenge if they are expected to achieve consensus in selecting an appropriate translation because words, phrases, jargon, or idiomatic expressions may cause flaws and erroneous conclusions about cultural differences. This suggests that no matter which methods are chosen, the translators influence the quality of the translation instrument. The high cost of the translation process, and the limited availability of expert translators, may pose added challenges (Carlson, 2000). Within the context of time, even if there is no set time limit for the translation, several stages of the translation process make it very tedious and time consuming. It is important to ensure that the time and cost involved in completing the translation are appropriately covered in the research proposal and budget (Carlson, 2000; Willgerodt et al., 2005). It is important that instrument developers take care to focus on the details and challenges of translation to avoid translation errors. Capitulo et al. (2001) identified four types of translation errors: type 1 errors occur when words or phrases are added to the original instrument; type 2 errors result from deleting words or phrases from the original instrument; type 3 errors are caused by altering words or phrases so that they no longer convey the same meaning of the original instrument; and type 4 errors stem from the use of poor grammar and syntax, which negatively affect the meaning and clarity.

Researchers (Capitulo et al., 2001; Eremenco et al., 2005; Duffy, 2006; Hilton & Skrutkowski, 2002) suggested ways to minimise translation errors which include

recruiting translators from native speakers who are able to read and write the target language fluently. Language experts who are educated in the language, and experts with formal preparation in translation, should also be recruited. Establishing formal standards and a policy for translation and maintaining readability level of grades 3–5 in the source and target languages would be helpful. Hall et al., (2003) also iterated that, the use of short and simple language will allow for an adequate and equivalent meaning of words and expressions. Utilising committee members to provide scientific translation services, could enhance the quality of the translation. The role of experts in all of the translation methods is important to capture the content and broader linguistic and cultural views (Eremenco et al., 2005; Ramirez et al., 2006). When circumstances or human and financial resources do not allow researchers to utilise back translation methods, other methods such as one way or expert’s translation or translation by a committee may be a better option. However, it is important that the researchers describe the translation procedure clearly and provide an indication of the quality of the translation when reporting their findings. It is also suggested that journal editors and reviewers be aware of the translation procedures to maintain high quality translations and proper reporting of the methods used by researchers (Peters & Passchier, 2006). Thus the translation process can be costly and labour intensive. However, “it is an essential step in generating research findings that can be deemed credible and valid by the scientific and professional community” (Capitulo et al., 2001:170).

CONCLUSION

The development of appropriate and valid multicultural and multilingual instruments in research is necessary because of an increasing multicultural and multilingual society in the 21st century. The general aim of the translation process should be to stay close to the original version of the instrument so that it is meaningful and easily understood by the target language population. Therefore, designing cross-cultural and meaningful research instruments requires the appropriate use of translation methodologies and testing processes. Translation methods and processes have many challenges, but applying relevant and creative strategies to reduce errors and pitfalls is crucial to achieve semantic equivalence.

REFERENCES

- Bowden, A. & Fox-Rushby, J.A. 2003. A systematic and critical review of the process of translation and adaptation of generic health-related quality of life measures in Africa, Asia, Eastern Europe, the Middle East, South Africa. *Social Science & Medicine*, 5:1289–1306.
- Capitulo, K.L., Cornelio, M.A. & Lenz, E.R. 2001. Clinical methods: translating the short version of the perinatal grief scale: process and challenges. *Applied Nursing Research*, 14(3): 165–170.
- Carlson, E.D. 2000. A case study in translation methodology using the health promotion lifestyle profile II. *Public Health Nursing*, 17(1):61–70.

- Duffy, M.E. 2006. Translating instruments into other languages. *Clinical Nurse Specialist*, 20(5):225–226.
- Eremenco, S.L., Cella, D. & Arnold, B.J. 2005. A comprehensive method for the translation and cross-cultural validation of health status questionnaires. *Evaluation Health Professionals*, 28(2):212–232.
- Hall, E.O.C., Wilson, M.E. & Frankenfield, J.A. 2003. Translation and restandardization of an instrument: the early infant temperament questionnaire. *Journal of Advanced Nursing*, 42(2):159–168.
- Hilton, A. & Skrutkowski, M. 2002. Translating instruments into other languages: developing and testing processes. *Cancer Nursing*, 25(1):1–7.
- Maneesriwongul, W. & Dixon, J.K. 2004. Instrument translation process: a methods review. *Journal of Advanced Nursing*, 48(2):175–186.
- Martinez, G., Marin, B.V., & Schoua-Glusberg, A. 2006. Translation from English to Spanish: the 2002 national survey of family growth. *Hispanic Journal of Behavioral Sciences*, 28(4): 531–545.
- National Child Traumatic Stress Network. 2006. Culture and trauma brief: translation of English materials to Spanish. *NCTSN Culture & Trauma Briefs*, 1(3). Available at: http://www.nctsn.org/nctsn_assets/pdfs/culture_and_trauma_brief_translations.pdf (accessed 05 October 2007).
- NCTSN – see National Child Traumatic Stress Network.
- Peters, M. & Passchier, J. 2006. Translating instruments for cross-cultural studies in headache research. *The Journal of Head & Face Pain*, 46(1):82–91.
- Polit, D.F., Beck, C.T., & Hungler, B.P. 2001. *Essentials of nursing research: methods, appraisal, and utilization*. 5th Edition. Philadelphia: J.B. Lippincott.
- Poss, J.E. 1999. Developing an instrument to study the tuberculosis screening behaviors of Mexican migrant farmworkers. *Journal of Transcultural Nursing*, 10(4):306–319.
- Ramirez, M., Teresi, J.A., Holmes, D., Gurland, B. & Lantigua, R. 2006. Differential item functioning (DIF) and the Mini-Mental State Examination (MMSE): overview, sample, and issues of translation. *Medical Care*, 44(11, Suppl. 3):S95–S105.
- Wang, W., Lee, H., & Fetzer, S.J. 2006. Challenges and strategies of instrument translation. *Western Journal of Nursing Research*, 28(3):310–321.
- Willgerodt, M.A., Kataoka-Yahiro, M., Kim, E. & Ceria, C. 2005. Issues of instrument translation in research on Asian immigrant populations. *Journal of Professional Nursing*, 21(4):231–239.