Methodological considerations when translating "burnout"☆

Allison Squires[a,*], **Catherine Finlayson**[a], **Lauren Gerchow**[a], **Jeannie P. Cimiotti**[b], **Anne Matthews**[c], **Rene Schwendimann**[d], **Peter Griffiths**[e], **Reinhard Busse**[f], **Maude Heinen**[g], **Tomasz Brzostek**[h], **Maria Teresa Moreno-Casbas**[i], **Linda H. Aiken**[j,k], and **Walter Sermeus**[k,l]

Catherine Finlayson: csf201@nyu.edu; Lauren Gerchow: lgerchow@gmail.com; Jeannie P. Cimiotti: cimiotti@rutgers.edu; Anne Matthews: anne.matthews@dcu.ie; Rene Schwendimann: rene.schwendimann@unibas.ch; Peter Griffiths: peter.griffiths@soton.ac.uk; Reinhard Busse: rbusse@tu-berlin.de; Maude Heinen: Maud.Heinen@Radboudumc.nl; Tomasz Brzostek: tomasz.brzostek@uj.edu.pl; Maria Teresa Moreno-Casbas: mmoreno@isciii.es; Linda H. Aiken: laiken@nursing.upenn.edu; Walter Sermeus: walter.sermeus@med.kuleuven.ac.be

[a]College of Nursing, New York University, USA [b]NJ Collaborating Center for Nursing, Rutgers University College of Nursing, USA [c]School of Nursing & Human Sciences, Dublin City University, Ireland [d]University of Basel, Institute of Nursing Science, Switzerland [e]Centre for Innovation and Leadership in Health Sciences, University of Southampton, England, United Kingdom [f]Lehrstuhl Management im Gesundheitswesen/Department of Health Care Management – WHO Collaborating Centre for Health Systems Research and Management, Technische Universitaet Berlin, Germany [g]IQ Healthcare, Radboud University Medical Center, Nursing Science & Allied Healthcare Division, Nijmegen, The Netherlands [h]Institute of Nursing and Midwifery, Faculty of Health Sciences, Jagiellonian University Medical College, Krakow, Poland [i]Nursing and Healthcare Research Unit (Investén-isciii), Spanish Department of Health, Madrid, Spain [j]School of Nursing, University of Pennsylvania, USA [k]RN4CAST, Spain [l]Program Director Master in Healthcare Management & Nursing Science Centre for Health Services & Nursing Research Catholic University Leuven, Belgium

Abstract

No study has systematically examined how researchers address cross-cultural adaptation of burnout. We conducted an integrative review to examine how researchers had adapted the

*Corresponding author at: College of Nursing, New York University, 726 Broadway, 10th Floor, New York, NY 10003, USA. Tel.: +1 212 992 7074. aps6@nyu.edu (A. Squires).

instruments to the different contexts. We reviewed the Content Validity Indexing scores for the Maslach Burnout Inventory-Human Services Survey from the 12-country comparative nursing workforce study, RN4CAST. In the integrative review, multiple issues related to translation were found in existing studies. In the cross-cultural instrument analysis, 7 out of 22 items on the instrument received an extremely low kappa score. Investigators may need to employ more rigorous cross-cultural adaptation methods when attempting to measure burnout.

Keywords

Europe; Nurses; Nursing; Burnout; Language translation; Content Validity Indexing; Human resources for health; Cross-cultural instrument adaptation

1. Introduction

Burnout is a global work-related phenomenon that, as multiple studies in mostly English-speaking developed countries have demonstrated, is associated with the quality of working conditions, interpersonal relationships, role conflict, and workload (Maslach, Schaufeli, & Leiter, 2001). Burnout is specifically a work-related syndrome that is commonly found in people who work in human services that require significant human contact (Maslach et al., 2001). For example, high levels of burnout have been reported among nurses in multiple countries (Aiken et al., 2012; Hatcher & Laschinger, 1996; McHugh, Kutney-Lee, Cimiotti, Sloane, & Aiken, 2011; Poghosyan, Clarke, Finlayson, & Aiken, 2010). Worker burnout, as a concept for research in the health professions, remains a cross-culturally relevant subject.

The Maslach Burnout Inventory is composed Likert-type items that assess three distinct components of burnout: emotional exhaustion, depersonalization, and personal accomplishment (Maslach et al., 2001). Three versions of the MBI exist, but the Maslach Burnout Inventory-Human Services Survey (MBI-HSS) has been designed exclusively for professionals whose work involves intensive human contact and interaction, such as nurses. Recent work continues to demonstrate that the MBI-HSS (Maslach et al., 2001) is the standard for gauging burnout among English-speaking healthcare professionals, especially nurses (Aiken, Clarke, Sloane, Sochalski, & Silber, 2002; Aiken et al., 2010; Cimiotti, Aiken, Sloane, & Wu, 2012; Hatcher & Laschinger, 1996; Laschinger, Grau, Finegan, & Wilk, 2010; Losa Iglesias, Becerro de Bengoa Vallejo, & Salvadores Fuentes, 2010; Patrician, Shang, & Lake, 2010; Poghosyan et al., 2010; Santen, Holt, Kemp, & Hemphill, 2010; Stimpfel, Sloane, & Aiken, 2012). Originally developed in English, the MBI-HSS has been used to assess burnout in nurses in several English-speaking countries, including the US (Aiken, Clarke, Sloane, & Sochalski, 2001; Aiken et al., 2002,2010), Canada (Aiken et al., 2001; Estabrooks et al., 2002; Hatcher & Laschinger, 1996; Laschinger et al., 2010; Leiter & Spence Laschinger, 2006; Spence Laschinger, & Leiter, 2006; Tourangeau et al., 2007), England (Sheward, Hunt, Hagen, Macleod, & Ball, 2005; Rafferty et al., 2007), Scotland (Aiken et al., 2001), and New Zealand (Poghosyan et al., 2010). The MBI-HSS has been translated into German (Aiken et al., 2001), Hebrew (Chayu & Kreitler, 2011), Japanese (Poghosyan et al., 2010), Turkish (Akku , Karacan, Göker, & Aksu, 2010; Günü en &Ustün, 2010) and Chinese (Yao, Yao, Wang, Li, & Lan, 2013).

Despite these translated uses of the MBI-HSS, reports of systematic translations of this instrument are few. The cross-cultural adaptation of any instrument designed in one country for use in another requires a rigorous and systematic process (Squires et al., 2013). For researchers in countries seeking to study burnout for the first time, both language and translatability may present problems to utilizing the MBI-HSS instrument itself. While studies in the literature have translated the MBI-HSS, authors have done so using a variety of techniques and analytic approaches. The variation in validation methods for translated versions of the MBI-HSS presents several methodological issues for researchers. One methodological challenge for studying burnout arises with the concept of "burnout" including the language used to express or define burnout and its dimensions, such as emotional exhaustion. These descriptions might not exist culturally or linguistically, or the general concept of burnout may be culturally taboo. For example, if the cultural norm for dealing with or expressing symptoms related to burnout is to endure them silently, the concept of burnout might not exist in the language of the culture or may remain unidentified or have a different descriptor. Yet we do know that it is a psychological syndrome that is commonly identified in people and thus, an initial study may be needed to identify the presence of the phenomenon in a new context or culture through a standardized measure. Comparatively, in cultures where emotional expressions of burnout are expected, burnout, or an equivalent concept may already exist in the language.

A good example of the cross-cultural and language translation challenges of burnout can present is how "burnout" as a concept has translated into Mexican Spanish. Balseiro Almario translated burnout as "fatiga emocional" and "desgaste emocional" when introducing this concept into the Mexican health services research literature (Balseiro-Almario, 2004, 2005). When translated back into English, the phrases respectively translate as "emotional fatigue" and "emotional exhaustion, which only describes one dimension of burnout. Despite coining these translated phrases, Balseiro Almario used the English word "burnout" in her articles, either for its ease of pronunciation in Spanish or because an adequate phrase in Spanish could not be found to encompass all the dimensions of burnout. The following questions then arise: (1) Why only focus the translation on one dimension of burnout? (2) What aspect of that phrasing makes the concept culturally presentable? These kinds of considerations and challenges are common when translating and applying complex concepts, like burnout, in new cultural and language contexts.

In the case of the RN4CAST (www.rn4cast.eu) study (Sermeus et al., 2011), in order to evaluate the effects of work environments on nurses' perceptions about their job related burdens, the study integrated the Maslach Burnout Inventory into a 12-country comparative study of nursing professionals in Europe. The initial review of the instrument by the RN4CAST research team raised concerns about the language used in the MBI-HSS and its translatability to other languages, cultures, countries, and contexts. It inspired the team to investigate the issue further.

The purpose this study was first to evaluate translated versions of the MBI-HSS that exist in the current literature to explore how other studies have conducted the cross-cultural adaptation of the MBI-HSS. The team then evaluated the MBI-HSS specific results of the translation process developed for the RN4CAST study instrument, which was grounded in

Flaherty et al.'s (1988) guidelines for cross-cultural evaluation of survey instruments. The systematic approach to translating the survey instrument included traditional forward and back translation techniques (Brislin, 1970), expert panel reviews, and a quantifiable technique to evaluate the relevance of questions and the quality of the translation, as described in detail by Squires et al. (2013). Overall, we seek to illustrate the challenges that can arise when translating complex concepts and to identify threats to the reliability and validity of study results when the translation of survey instruments is not conducted systematically and without consideration for the cross-cultural relevance of the topic.

1.1. Background

The World Health Report 2006 – Working Together for Health (World Health Organization, 2006) is an assessment of the global healthcare workforce that estimates a shortage of 4.3 healthcare personnel per 100,000 people worldwide. The report further cites job-related burnout and its associated outcomes as factors that contribute to healthcare workers intentions to leave their jobs. Healthcare worker burnout appears to be an economic burden to national health systems and its care organizations that are struggling to maintain adequate staff numbers to provide health services to those in need (Alameddine, Baumann, Laporte, & Deber, 2012; El-Jardali et al., 2011; Pomaki, Franche, Murray, Khushrushahi, & Lampinen, 2012; Stansfeld, Shipley, Head, & Fuhrer, 2012). It is beneficial for healthcare systems to evaluate levels of burnout among their staff to determine its effects on retention rates, attrition rates from the profession, or its effects on international migration (Alameddine et al., 2012).

The MBI-HSS has been the chosen instrument of many researchers looking to gain better insight into the problem of burnout among healthcare professionals. While many studies have used the MBI-HSS to measure burnout across different cultures and national contexts, our review of the literature for this study showed that translation methods and study designs appear inconsistent. Researchers with extensive cross-cultural research experience generally agree that a systematic approach is necessary when translating a concept like burnout. A solid translation process will manage the emic and etic aspects of translation, include forward and back translation completed through the use of qualified translators, and an expert panel review (Cha, Kim, & Erlen, 2007; Hilton & Skrutkowski, 2002; Hyrkäs, Appelqvist-Schmidlechner, & Oksa, 2003; Im, Page, Lin, Tsai, & Cheng, 2004; Jones, Lee, Phillips, Zhang, & Jaceldo, 2001; Sidani, Guruge, Miranda, Ford-Gilboe, & Varcoe, 2010; Temple, 2005; Wang, Lee, & Fetzer, 2006; Weeks, Swerissen, & Belfrage, 2007). In addition, a pilot study of the translated version of the instrument is also recommended.

More concerning, however, is that health researchers tend to only use forward and back translation techniques. Maneesriwongul and Dixon (2004) showed, through a rigorous integrative review of studies that required translating survey instruments, that simple forward and backward translation is insufficient to produce a valid translation. Therefore, the results from studies that only employ forward and backward translation of survey instruments for their translation approach might not be very reliable or valid because of the quality of the translation or produce artificially high or low results.

In light of the concerns about translation quality, when examining the MBI-HSS prior to study implementation, the RN4CAST research team found that the survey items include questions that use slang words or phrases in American English. Unless researchers conducting burnout studies were familiar with the appropriate meaning and interpretation of the slang words, or they used an American English speaker to conduct the translation, the validity of the translation and the respective survey findings could be called into question. Who conducted the translation can be an important factor in evaluating the quality of the translation and its threat to reliability and validity (Squires, 2008,2009; Temple & Young, 2004; Temple, 2002).

An evaluation of the cross-cultural relevance of an established instrument, which includes translation quality, is an important step prior to its use outside of its original context of development. While it may not be possible to alter the instrument when consistent results across countries are part of a study's aims, a pre-data collection evaluation step may help researchers to identify potentially problematic items in the instrument that are specific to the setting, language, or culture (Squires et al., 2013). At the same time, if a valid translation exists, duplicating work wastes time.

2. Methods and results

We explored the issue of cross-cultural adaptation of the MBI-HSS in two ways. First, we conducted an integrative literature review focused on translation methods used in research studies that had translated the MBI-HSS, and second, a Content Validity Indexing (CVI) process to evaluate the cross-cultural relevance of the translations conducted by the study's country teams. We present the approaches and results for both parts of the study in two sections and then synthesize the findings in the discussion.

2.1. Approach to the integrative review of the literature

For the integrative review, the team used MEDLINE, CINAHL and Google Scholar, and searched for articles about the MBI that involved language translation. Search terms included "Maslach", "burnout", "international," "language," "interpretation," and "translation". Over 400 articles were identified in the search process. Selection criteria for articles included their availability in English, Spanish, or Portuguese, publication date after 2000.

Once the team finalized article selection, directed content analysis techniques, defined as "a research method for the subjective interpretation of the content of text data through the systematic classification process of coding and identifying themes or patterns" (Hsieh & Shannon, 2005), provided structure to the review process. To code the articles, the team focused on identifying the presence of cross-cultural research characteristics, as defined by Flaherty et al. (1988) and illustrated in Table 1, the translation method used, and when applicable, the type of statistical analysis used to determine the reliability and validity of the translated instrument. Detailed notes about specific methodological issues related to translation or how researchers interpreted the results were also made during the coding process.

2.2. Results of the integrative review

In the review, 30 articles met the criteria for the analysis, representing 26 countries, 20 languages, and 8 of the RN4CAST countries. There were 7 languages with regional or geographical dialects represented, including Spanish (Spain, Argentina, Colombia), German (Germany and Switzerland), Italian (Italy and Switzerland), French (France and Switzerland), Dutch (Belgium and the Netherlands), Chinese (Cantonese, Mandarin, Macau dialect, and Taiwanese), and Arabic (Yemeni dialect).

Table 2 summarizes the findings, including the cross-cultural analysis against Flaherty et al.'s (1988) criteria. Within the 30 articles analyzed, there were vast differences in each approach to translate the MBI-HSS. Literature analyzed for this study offered no consistent findings about the cross-cultural relevance of burnout and showed that researchers used no consistent approach to translating the instrument, even when a translated version already existed in their own language.

Among the 30 reviewed articles, 7 (22%) articles did not mention any method of translation and 9 (28%) articles cited other validated versions of MBI-HSS in their respective language. The majority of the rest used only forward and backward translation. A few used a more thorough cross-cultural validation process, usually involving an expert panel review or a pilot study.

The 30 articles were further analyzed against five criteria suggested by Flaherty, et al. (Flaherty et al., 1988): content equivalence, semantic equivalence, technical equivalence, criterion equivalence, and conceptual equivalence. Every article met the criteria for content equivalence, which determined if the MBI-HSS was the appropriate tool for their study, and aimed to determine if burnout was present in the study population. Since the MBI-HSS was administered via questionnaire in all 30 articles, they met the technical equivalence requirement, which determined if the tool was being administered appropriately in the context.

Semantic equivalence was more difficult to determine. Semantic equivalence is defined by Flaherty et al. as "the meaning of each item is the same in each culture after translation into the language and idiom (written or oral) of each cultures" (Flaherty et al., 1988). A tool could be deemed semantically equivalent if it has been validated through a rigorous process. However, many of the articles did not mention how the translation was conducted nor indicated if they used a previously validated version of the MBI. Therefore, the team concluded that many of these articles did not meet the criteria for semantic equivalence. None of the articles met the definition for criterion equivalence, which evaluates if "the interpretation of the measurement of the variable remains the same when compared with the norm for each culture studied" (Flaherty et al., 1988). None of the articles met this criterion because it was not apparent if any of the studies met the norm for the culture being studied. Finally, it was nearly impossible to determine how authors of the studies addressed conceptual equivalence, which would measure "the same theoretical construct in each culture" (Flaherty et al., 1988).

2.3. RN4CAST translation approach and cross-cultural evaluation

For this paper, we examined the study's pre-data collection evaluations of the cross-cultural relevance of the translated versions of the MBI-HSS survey questions produced by the RN4CAST country team members. This included eleven European countries with ten languages among them. They included Belgium (Dutch and French), Germany (German), Finland (Finnish), Greece (Greek), Ireland (Irish-English), the Netherlands (Dutch), Poland (Polish), Spain (Spanish), Sweden (Swedish), Switzerland (Swiss French, Swiss German, Swiss Italian), and England (British English). For baseline comparison, data from the United States (American English) were also included. At the time the study began, a valid Spanish translation approved by the company that owns the copyright of the MBI-HSS was available so Spain was the only country that used an official translation, and they did not participate in this pre-data collection evaluation exercise.

Prior to beginning the study, the RN4CAST team concluded that it could not use any of the existing versions of the translated instrument. Therefore, with the permission of the company that manages the MBI-HSS, the team conducted its own approach to translating and evaluating the cross-cultural relevance of the instrument. The translation process involved multiple steps and met the requirements for rigorous cross-cultural translation of established instruments, including an evaluation of content, context, conceptual, semantic, and technical equivalence. It involved two phases of reviews that included nurses with extensive experience in their field, the use of experienced translators who used a translation guide to facilitate translation, and a quantification of relevance of survey items to nursing in the country. By evaluating the "relevance" of survey items, we mean that the "expert" nurses reviewing the survey item determine if the question is appropriate for use with the population to be studied and in the case of this study, in the cultural context(s) of the country where is would be applied. We chose nurses as our study experts rather than occupational psychologists or psychiatrists to ensure that the translated instrument was validated by those intimately involved with the population who would be surveyed with the MBI-HSS. Their content and contextual knowledge were essential components of the expertise we sought in "expert" raters. Complete details of the approach to translation are found in Squires et al. (2013).

In order to produce the quantifiable measure of cross-cultural relevance and translation quality, the research team opted to use an approach normally advocated for initial instrument development: Content Validity Indexing (CVI) with corrections for chance agreement. During the CVI process, expert raters provide scored feedback to determine if the question or statement, e.g., in a survey instrument, is relevant to the population being studied and if the format of the question is appropriate (Polit, Beck, & Owen, 2007). This assessment of a question's "relevance" allows an expert to determine the content, context, and conceptual equivalence of a translated survey question through a single rating. The experts used the following standard CVI rating scale to evaluate a question's relevance: 1 = Not relevant; 2 = Somewhat relevant; 3 = Very relevant; 4 = Highly relevant. A calculation of chance correction, via a modified kappa score, adjusts the CVI score to indicate agreement over and above chance, thereby increasing the rigor of the CVI process (Polit et al, 2007). Only scores of 3 or 4 are included in the calculation process.

The CVI cannot evaluate the semantic and technical equivalence of a translated question in its rating process. Therefore, the team added to the CVI process a simple "yes" or "no" rating to evaluate the semantic and technical equivalence of the translations. It was the team's intention to keep this aspect of the rating process as simple as possible to bolster confidence in the cross-cultural translation process. Comments sections for raters allow for recommended corrections to translations, as needed.

Both scoring systems use the same formula for calculating item-level and scale-level scores with modified kappa statistics for comparison. For the CVI process, I-CVI represents the item level score where the minimum acceptable result is 0.78 or higher (without correction for chance agreement among raters) to include an item in a survey. The modified kappa score equivalent to the 0.78 CVI score is 0.74, but offers a comparison of scores that would be considered "good" (0.60–0.73) vs. "excellent" (0.74 or higher) (Polit et al., 2007). A modified kappa score of 0.59 or lower would mean the item was not acceptable for inclusion in the survey or that the content or phrasing of the question might need modification (Squires et al., 2013). The scale level CVI score (S-CVI) is an average of all ratings for all items in the instrument.

The team then had to agree on how to manage low scores from raters on individual items. Since the MBI-HSS is an established instrument, the team could not discard any items from the survey. Therefore, the team agreed to set minimum levels for scores in order to "approve" the translation for use in the larger study. Survey questions that received low scores from the raters, a modified kappa score of 0.59 or lower, were identified as potentially problematic items. Each team would note the potentially problematic items and then work to aggregate the pre-data collection data to modify the translation when needed. In the larger study, results keyed to potentially problematic items receive closer scrutiny for trends that might suggest an issue with cross-cultural equivalence.

A total of 120 raters from the twelve countries involved were invited to participate. Expert raters from each country completed the evaluation process through an online survey. Raters were asked to report their educational level and current role: Practitioner (meaning working in a front line patient care role), educator (involved in health education in a hospital or educational institution), or administrator (hospital or other organization). Upon completion of the rating process, the survey company emailed the results to the project manager in a spreadsheet format. Preprogrammed formulas in Microsoft Excel calculated the item and survey level data using the CVI with chance correction method described by Polit et al. (2007).

2.4. Results of the RN4CAST cross-cultural CVI analysis

A total of 106 raters out of 120 participated in the CVI with the translation evaluation process and their basic demographics are found in Table 3. With one exception, Germany (*n* = 5), each country had 7 or more raters with a maximum of 11. Germany was only able to obtain five raters to participate in the process because there was a significant lack of bilingual German-English speaking nurses in the country. Complete results of the MBI-HSS CVI rating process with chance correction are found in Table 4, illustrating the S-CVI scores. Generally, the scale-level MBI-HSS scores varied widely and ranged from 0.49 (US

and Belgian French) to 0.93 (Germany). Out of the ten languages and their country variants involved in the validation process, half received acceptable scale level scores, while half did not. Fig. 1 illustrates 5 out of 7 items on the MBI-HSS that received extremely low scores by 60% or more of participating countries and these items were identified as likely to produce "problematic" responses in the larger RN4CAST study.

Table 5 illustrates the problematic item rating trends across the languages and regions involved in the study using 5 items from Fig. 1. Similar languages are placed side-by-side in the table for comparison. While the sample size limited our ability to test for significant differences between the groups, some trends are evident and are worthy of analysis. To begin, Greece and the Italian region of Switzerland had very high relevance scores with all items on the MBI-HSS, suggesting that the MBI-HSS could be a useful tool in assessing burnout for nurses or that expressing burnout is culturally acceptable in these regions. Some countries that spoke the same language (such as Belgian French and Swiss French) gave very similar individual question scores. Swiss German and German, on the contrary, were exactly opposite in their ratings of all questions. Belgian Dutch and the Netherlands' Dutch were mixed.

3. Discussion

The notable lack of attention paid to translation processes in the studies evaluated for the integrative review and the variability of the results and our own subsequent evaluations suggest that researchers and journal editors should pay more attention to how authors have conducted translation processes for the MBI-HSS. Statistical analyses of survey results to analyze their validity will not capture problems related to translation, which occurs before data collection. Our findings also suggest that outliers in study results may be the result of an issue with the translation process. Factor loading on factor analyses may also deviate and affect subscale composition and thus, the consistency of the scale across contexts and cultures.

Some I-CVI scores on the MBI-HSS were lower than anticipated. A possible explanation for the lower MBI-HSS scores, in the case of this paper, concerns whether or not the concept of or language for "burnout" is actually present among the nursing workforce vernacular in a country. Nurses may very well possess the signs and symptoms of burnout, but they may not yet have a name for what they are experiencing and may only report high levels of emotional exhaustion. For some cultures, even the idea of feeling "burnt out" in a job may not align with their cultural norms and values; therefore, nurses and other healthcare workers might ignore or suppress symptoms of burnout.

The remainder of the discussion proceeds on a country-by-country or language "case" basis. To begin, several explanations might shed light on the results from the English-languages versions rating process. The first concerns the use of American English slang in the MBI-HSS. Comments from the Irish and English raters suggested that while they understood the intent of the question, the language describing the concept was not expressed in a way that was common in their home country. This underscores the need to cross-culturally validate even English language instruments when used in another country.

English speaking countries also scored the MBI-HSS relatively low on the relevance of "burnout" to nursing practice in their home countries. This was a surprising finding for the team. One explanation for these scores included expert sampling bias, where "experts" worked in organizations with supportive organizational cultures for nursing practice. Therefore, they did not perceive the "burnout" questions as relevant to practice compared to a nurse who might work in a less supportive organization. System-wide reforms in both English speaking countries might also contribute to changed perceptions about the relevance of burnout. Language and phrasing of the items may also have differed sufficiently for the experts to score the items as less relevant. The number of internationally educated nurses in both the UK and Ireland results in a variety of "other Englishes" as well, thereby presenting additional challenges of harmonizing English language phrasing between healthcare workers from different countries that speak English. Future research may need to better account for issues of rater identity during this process.

In the case of Poland, many questions had modified kappa scores below 0.60. Consequently, the RN4CAST team recommended that the Polish team pilot test the entire instrument before using it in the larger study (Brzostek et al., in press). The relative "newness" of nursing research in Poland may also explain some of the low CVI scores from that country. The Polish case highlights the importance of a rigorous cross-cultural adaptation process when using the MBI-HSS in a new country.

The breadth of scores in Switzerland is puzzling to the outside researcher. The Swiss team, however, was not surprised by the variation among the scores between the three linguistic regions of the country as they commonly find variations in research results when conducting national workforce studies. The Swiss team, therefore, made note of the potentially problematic items from their rating process (beyond the 7 identified by the entire study). They concluded that if outlier responses in the Swiss burnout results were noted a problem with the contextual applicability of the item might be indicated. The Belgian team adopted a similar tactic with the Belgian French version of the MBI-HSS to address its low score. The findings from both countries suggest that the linguistic expression of burnout has a cultural component that researchers need to anticipate in comparative national studies of healthcare workers and within country comparisons when more than one official language exists. Additional analyses to explore this phenomenon are planned.

3.1. Limitations

With twelve countries and ten languages, limitations for this study relate mostly to the individuals conducting the CVI rating process. Germany, notably, had only five raters, which skewed its results toward higher scores even though comments provided by raters suggested some issues related to the translation of terms found in the seven problematic items. It is possible that with more raters their results may have looked more like the Swiss-German version. Additionally, language abilities and educational levels of the raters varied across the countries and were not always formally verifiable through English language test scores or other measures of language competence. This likely affected scoring processes due to widely varying bilingual capabilities among nurses. The use of only raters with nursing experience may have also biased the results. We recommend that researchers seeking to

replicate these methods find a way to account for the linguistic capabilities of their expert raters. For the integrative review, due to our linguistic limitations for the search results and the possibility that a national journal (which might contain a good translation of the MBI-HSS) was not listed in the search databases, we might not have been able to capture all of the available studies.

3.2. Conclusions

To the best of our knowledge this is the first time a mixed-methods evaluation of the translatability of the MBI-HSS occurred through a structured research approach. The RN4CAST team plans to conduct additional analyses confirming the validity and reliability of the instrument across the different national contexts. Results of this study have lead us to encourage individual country-level validation processes for research teams seeking to use translated instruments for research projects, even if the country is linguistically similar the instrument's country of origin.

We also note that as health services research spreads across the globe, researchers using survey instruments in their studies should be cautioned to carefully and systematically translate and evaluate the cross cultural relevance of the instrument prior to data collection. Our literature reviews shows that this is often not done. Most survey instruments were developed for use in the country of origin of the researcher who created it. There is no guarantee that the concepts measured by the instrument will apply in the same way in another country and language context, especially when health systems are organized in different ways. Studies that fail to pre-evaluate the cross-cultural applicability of a survey instrument may not produce reliable and valid results. Using systematic approaches can help to reduce the risk for those kinds of errors.

Acknowledgments

References

Aiken LH, Clarke SP, Sloane DM, Sochalski J. Cause for concern: Nurses' reports of hospital care in five countries. LDI Issue Brief. 2001; 6(8):1–4. Available from: http://www.ncbi.nlm.nili.gov/pubmed/12524707. [PubMed: 12524707]

Aiken LH, Clarke SP, Sloane DM, Sochalski J, Silber JH. Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction. Journal of the American Medical Association. 2002; 288(16):1987–1993. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12387650. [PubMed: 12387650]

Aiken LH, Sloane DM, Cimiotti JP, Clarke SP, Flynn L, Seago JA, et al. Implications of the California nurse staffing mandate for other states. Health Services Research. 2010; 45(4):904–921. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=200&tool=pmcentrez&rendertype=Abstract2908. [PubMed: 20403061]

Aiken LH, Sermeus W, Van den Heede K, Sloane DM, Busse R, McKee M, et al. Patient safety, satisfaction, and quality of hospital care: Cross sectional surveys of nurses and patients in 12

countries in Europe and the United States. British Medical Journal. 2012; 344:e1717. Available from: http://www.bmj.com/cgi/doi/10.1136/bmj.e1717. [PubMed: 22434089]

Akku Y, Karacan Y, Göker H, Aksu S. Determination of burnout levels of nurses working in stem cell transplantation units in Turkey. Nursing & Health Sciences. 2010; 12(4):444–449. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21210922. [PubMed: 21210922]

Alameddine M, Baumann A, Laporte A, Deber R. A narrative review on the effect of economic downturns on the nursing labour market: Implications for policy and planning. Human Resources for Health. 2012; 10(1):23. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=285&tool=pmcentrez&rendertype=Abstract3472. [PubMed: 22905739]

Balseiro-Almario L. El sindrom de fatiga emocional (burnout): Una vieja patologia de enfermería, que debe ser atendida en forma urgente [Burnout: An old nursing pathology that needs to be immediately addressed]. Enfermería Universitaria. 2004; 1(1):21–23.

Balseiro-Almario L. El síndrome de burnout: Un riesgo permanente en el trabajo de enfermería [Burnout syndrome: A permanent risk in nursing]. Enfermería Universitaria. 2005; 2(3):28–31.

Brislin RW. Back-translation for cross-cultural research. Journal of Cross-Cultural Psychology. 1970; 1(3):185–216. Available from: http://jcc.sagepub.com/content/l/3/185.short2014.

Cha E-S, Kim KH, Erlen JA. Translation of scales in cross-cultural research: Issues and techniques. Journal of Advanced Nursing. 2007; 58(4):386–395. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17442038. [PubMed: 17442038]

Chayu T, Kreitler S. Burnout in nephrology nurses in Israel. Nephrology Nursing Journal. 2011; 38(1):65–77. quiz 78. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21469556. [PubMed: 21469556]

Cimiotti JP, Aiken LH, Sloane DM, Wu ES. Nurse staffing, burnout, and health care-associated infection. American Journal of Infection Control. 2012; 40(6):486–490. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=207&tool=pmcentrez&rendertype=Abstract3509. [PubMed: 22854376]

El-Jardali F, Alameddine M, Dumit N, Dimassi H, Jamal D, Maalouf S. Nurses' work environment and intent to leave in Lebanese hospitals: Implications for policy and practice. International Journal of Nursing Studies. 2011; 48(2):204–214. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20932524. [PubMed: 20932524]

Estabrooks CA, Tourangeau AE, Humphrey CK, Hesketh KL, Giovannetti P, Thomson D, et al. Measuring the hospital practice environment: A Canadian context. Research in Nursing & Health. 2002; 25(4):256–268. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12124720. [PubMed: 12124720]

Flaherty JA, Gaviria FM, Pathak D, Mitchell T, Wintrob R, Richman JA, et al. Developing instruments for cross-cultural psychiatric research. Journal of Nervous and Mental Disease. 1988; 176(5):257–263. [PubMed: 3367140]

Günü en NP, Ustün B. An RCT of coping and support groups to reduce burnout among nurses. International Nursing Review. 2010; 57(4):485–492. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21050201. [PubMed: 21050201]

Hatcher S, Laschinger HK. Staff nurses' perceptions of job empowerment and level of burnout: A test of Kanter's theory of structural power in organizations. Canadian Journal of Nursing Administration. 1996; 9(2):74–94. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8716472. [PubMed: 8716472]

Hilton A, Skrutkowski M. Translating instruments into other languages: Development and testing processes. Cancer Nursing. 2002; 25(1):1–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11838715. [PubMed: 11838715]

Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qualitative Health Research. 2005; 15(9):1277–1288. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16204405. [PubMed: 16204405]

Hyrkäs K, Appelqvist-Schmidlechner K, Oksa L. Validating an instrument for clinical supervision using an expert panel. International Journal of Nursing Studies. 2003; 40(6):619–625. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0020748903000361. [PubMed: 12834927]

Im E-O, Page R, Lin L-C, Tsai HM, Cheng C-Y. Rigor in cross-cultural nursing research. International Journal of Nursing Studies. 2004; 41(8):891–899. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15476762. [PubMed: 15476762]

Jones PS, Lee JW, Phillips LR, Zhang XE, Jaceldo KB. An adaptation of Brislin's translation model for cross-cultural research. Nursing Research. 2001; 50(5):300–304. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11570715. [PubMed: 11570715]

Laschinger HKS, Grau AL, Finegan J, Wilk P. New graduate nurses' experiences of bullying and burnout in hospital settings. Journal of Advanced Nursing. 2010; 66(12):2732–2742. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20722806. [PubMed: 20722806]

Leiter MP, Spence Laschinger HK. Relationships of work and practice environment to professional burnout. Nursing Research. 2006; 55(2):137–146. Available from: http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage&an=00006199-200603000-00009. [PubMed: 16601626]

Losa Iglesias ME, Becerro de Bengoa Vallejo R, Salvadores Fuentes P. The relationship between experiential avoidance and burnout syndrome in critical care nurses: A cross-sectional questionnaire survey. International Journal of Nursing Studies. 2010; 47(1):30–37. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19625023. [PubMed: 19625023]

Maneesriwongul W, Dixon JK. Instrument translation process: A methods review. Journal of Advanced Nursing. 2004; 48(2):175–186. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15369498. [PubMed: 15369498]

Maslach C, Schaufeli WB, Leiter MP. Job burnout. Annual Review of Psychology. 2001; 52:397–422. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11148311.

McHugh MD, Kutney-Lee A, Cimiotti JP, Sloane DM, Aiken LH. Nurses' widespread job dissatisfaction, burnout, and frustration with health benefits signal problems for patient care. Health Affairs (Project Hope). 2011; 30(2):202–210. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3201822&tool=pmcentrez&rendertype=Abstract. [PubMed: 21289340]

Patrician PA, Shang J, Lake ET. Organizational determinants of work outcomes and quality care ratings among Army Medical Department registered nurses. Research in Nursing & Health. 2010; 33(2):99–110. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=846&tool=pmcentrez&rendertype=Abstract2969. [PubMed: 20151409]

Poghosyan L, Clarke SP, Finlayson M, Aiken LH. Nurse burnout and quality of care: Cross-national investigation in six countries. Research in Nursing & Health. 2010; 33(4):288–298. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=908&tool=pmcentrez&rendertype=Abstract2908. [PubMed: 20645421]

Polit DF, Beck T, Owen SV. Focus on research methods: Is the CVI an acceptable indicator of content validity? Appraisal and Recommendations. Research in Nursing & Health. 2007; 30:459–467. [PubMed: 17654487]

Pomaki G, Franche R-L, Murray E, Khushrushahi N, Lampinen TM. Workplace-based work disability prevention interventions for workers with common mental health conditions: A review of the literature. Journal of Occupational Rehabilitation. 2012; 22(2):182–195. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22038297. [PubMed: 22038297]

Rafferty AM, Clarke SP, Coles J, Ball J, James P, McKee M, et al. Outcomes of variation in hospital nurse staffing in English hospitals: Cross-sectional analysis of survey data and discharge records. International Journal of Nursing Studies. 2007; 44(2):175–182. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=580&tool=pmcentrez&rendertype=Abstract2894. [PubMed: 17064706]

Santen SA, Holt DB, Kemp JD, Hemphill RR. Burnout in medical students: Examining the prevalence and associated factors. Southern Medical Journal. 2010; 103(8):758–763. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20622724. [PubMed: 20622724]

Sermeus W, Aiken LH, Van den Heede K, Rafferty AM, Griffiths P, Moreno-Casbas MT, et al. Nurse forecasting in Europe (RN4CAST): Rationale, design and methodology. BMC Nursing. 2011; 10:6. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=324&tool=pmcentrez&rendertype=Abstract3108. [PubMed: 21501487]

Sheward L, Hunt J, Hagen S, Macleod M, Ball J. The relationship between UK hospital nurse staffing and emotional exhaustion and job dissatisfaction. Journal of Nursing Management. 2005; 13(1): 51–60. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15613094. [PubMed: 15613094]

Sidani S, Guruge S, Miranda J, Ford-Gilboe M, Varcoe C. Cultural adaptation and translation of measures: An integrated method. Research in Nursing & Health. 2010; 33(2):133–143. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20069612. [PubMed: 20069612]

Spence Laschinger HK, Leiter MP. The impact of nursing work environments on patient safety outcomes: The mediating role of burnout/engagement. Journal of Nursing Administration. 2006; 36(5):259–267. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16705307. [PubMed: 16705307]

Squires A. Language barriers and qualitative nursing research: Methodological considerations. International Nursing Review. 2008; 55(3):265–273. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=452&tool=pmcentrez&rendertype=Abstract2697. [PubMed: 19522941]

Squires A. Methodological challenges in cross-language qualitative research: A research review. International Journal of Nursing Studies. 2009; 46(2):277–287. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=094&tool=pmcentrez&rendertype=Abstract2784. [PubMed: 18789799]

Squires A, Aiken LH, Van den Heede K, Sermeus W, Bruyneel L, Lindqvist R, et al. A systematic survey instrument translation process for multi-country, comparative health workforce studies. International Journal of Nursing Studies. 2013; 50(2):264–273. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22445444. [PubMed: 22445444]

Stansfeld SA, Shipley MJ, Head J, Fuhrer R. Repeated job strain and the risk of depression: Longitudinal analyses from the Whitehall II study. American Journal of Public Health. 2012; 102(12):2360–2366. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=314&tool=pmcentrez&rendertype=Abstract3519. [PubMed: 23078508]

Stimpfel AW, Sloane DM, Aiken LH. The longer the shifts for hospital nurses, the higher the levels of burnout and patient dissatisfaction. Health Affairs (Project Hope). 2012; 31(11):2501–2509. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=421&tool=pmcentrez&rendertype=Abstract. [PubMed: 23129681]

Temple B. Crossed wires: Interpreters, translators, and bilingual workers in cross-language research. Qualitative Health Research. 2002; 12(6):844–854. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12109728. [PubMed: 12109728]

Temple B. Nice and tidy: Translation and representation. Sociological Research Online. 2005; 10(2) Available from: http://www.socresonline.org.Uk/10/2/temple.html.2012.

Temple B, Young A. Qualitative research and translation dilemmas. Qualitative Research. 2004; 4(2): 161–178. Available from: http://qrj.sagepub.com/content/4/2/161.short2014.

Tourangeau AE, Doran DM, McGillis Hall L, O'Brien Pallas L, Pringle D, Tu JV, et al. Impact of hospital nursing care on 30-day mortality for acute medical patients. Journal of Advanced Nursing. 2007; 57(1):32–44. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17184372. [PubMed: 17184372]

Wang WL, Lee HL, Fetzer SJ. Challenges and strategies of instrument translation. Western Journal of Nursing Research. 2006; 28(3):310–321. Available from: http://www.ncbi.nlm.nili.gov/pubmeci/16585807. [PubMed: 16585807]

Weeks A, Swerissen H, Belfrage J. Issues, challenges, and solutions in translating study instruments. Evaluation Review. 2007; 31(2):153–165. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17356181. [PubMed: 17356181]

World Health Organization. Working together for health: The world health report 2006. Geneva: WHO; 2006. Available from: http://www.who.int/whr/2006/whr06_en.pdf?ua=1

Yao Y, Yao W, Wang W, Li H, Lan Y. Investigation of risk factors of psychological acceptance and burnout syndrome among nurses in China. International Journal of Nursing Practice. 2013; 19(5): 530–538. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24093745. [PubMed: 24093745]

References included in the integrative review

Ahola K, et al. Contribution of burnout to the association between job strain and depression: The health 2000 study. Journal of Occupational and Environmental Medicine. 2006; 48(10):1023–1030. [PubMed: 17033502]

Al-Dubai SAR, Rampal KG. Prevalence and associated factors of burnout among doctors in Yemen. Journal of Occupational Health. 2010; 52(1):58–65. [PubMed: 19907108]

Asai M, et al. Burnout and psychiatric morbidity among physicians engaged in end-of-life care for cancer patients: A cross-sectional nationwide survey in Japan. Psycho-Oncology. 2007; 16(5):421–428. [PubMed: 16929464]

Berg AM, Hem E, Lau B, Ekeberg Ø. An exploration of job stress and health in Norwegian police service: A cross sectional study. Journal of Occupational Medicine and Toxicology. 2005; 55(2):113–120.

Bressi C, et al. Burnout among psychiatrists in Milan: A multicenter survey. Psychiatric Services. 2009; 60(7):985–989. [PubMed: 19564233]

Chen WS, Haniff J, Siau CS, Seet W, Loh SF, Abd Jamil MH. Pilot study of the Malay Maslach burnout inventory and Malay work related quality of life scale in Malaysia. Studies in Asian Social Science. 2013:20. http://dx.doi.org/10.5430/sass.v1n1.

Córdoba L, Tamayo JA, Gonzálex MA, Martínez MI, Rosales A, Barbato SH. Adaptation and validation of the Maslach Burnout Inventory-Human Services Survey in Cali, Columbia. Columbia Médico. 2011; 42:286–293.

Embriaco N, et al. High level of burnout in intensivists. American Journal of Respiratory and Critical Care Medicine. 2006; 175(7):686–692. [PubMed: 17234905]

Embriaco N, Papazian L, Kentish-Barnes N, Pochard F, Azoulay E. Burnout syndrome among critical care healthcare workers. Current Opinion in Critical Care. 2007; 13(5):482–488. [PubMed: 17762223]

Glasberg AL, Eriksson S, Norberg A. Burnout and 'stress of conscience' among healthcare personnel. Journal of Advanced Nursing. 2007; 57(4):392–403. [PubMed: 17291203]

Goehring G, Gallacchi MB, Künzi B, Bovier P. Psychosocial and professional characteristics of burnout in Swiss primary care practitioners: A cross-sectional survey. Swiss Medical Weekly. 2005; 135:101–108. [PubMed: 15832226]

Hu Q, Schaufeli WB. The factorial validity of the Maslach burnout inventory student survey in china. Psychological Reports. 2009; 105(2):394–408. [PubMed: 19928601]

Iglesias MEL, Vallejo RBB, Fuentes PS. The relationship between experiential avoidance and burnout syndrome in critical care nurses: A cross-sectional questionnaire survey. International Journal of Nursing Studies. 2010; 47(1):30–37. [PubMed: 19625023]

Juthberg C, Eriksson S, Norberg A, Sundin K. Perceptions of conscience, stress of conscience and burnout among nursing staff in residential elder care. Journal of Advanced Nursing. 2010; 66(8):1708–1718. [PubMed: 20557396]

Kanste O, Miettunen J, Kyngäs H. Factor structure of the Maslach burnout inventory among Finnish nursing staff. Nursing and Health Sciences. 2006; 8(4):201–207. [PubMed: 17081145]

Klersy C, et al. Burnout in health care providers of dialysis service in northern Italy - A multicentre study. Nephrology Dialysis Transplantation. 2007; 22(8):2283–2290.

Lee HF, Chein TW, Yen M. Examining factor structure of Maslach burnout inventory among nurses in Taiwan. journal of Nursing Management. 2013; 21:648–656. [PubMed: 23410056]

Liakopoulou M, et al. Burnout, staff support, and coping in pediatric oncology. Supportive Care in Cancer. 2008; 16(2):143–150. [PubMed: 17619909]

Luk AL, Chan BPS, Cheong SW, Ko SKK. An exploration of the burnout situation on teachers in two schools in Macau. Social Indicators Research. 2010; 95(3):489–502. [PubMed: 20062816]

Mészáros V, Ádám Sz, Szabó M, Szigeti M, Urbán R. The bifactor model of the Maslach Burnout Inventory-Human Services Survey – An alternative measurement model of burnout. Stress Health. 2013:2481. http://dx.doi.org/10.1002/smi.

Ndetei DM, et al. Burnout in staff working at the Mathari psychiatric hospital. African Journal of Psychiatry. 2008; 11(3):199–203. [PubMed: 19588043]

Pisanti R, Lombardo C, Lucidi F, Violani C, Lazzai D. Psychometric properties of the Maslach burnout inventory for human services among Italian nurses: A test of alternative models. Journal of Advanced Nursing. 2012; 69(3):697–707. [PubMed: 22897490]

Schaufeli WB, Bakker AB, van der Heijden FMMA, Prins JT. Workaholism, burnout and well-being among junior doctors: The mediating role of role conflict. Work and Stress. 2009; 23(2):155–172.

Tokuda Y, Hayano K, Ozaki M, Bito S, Yanai H, Koizumi S. The interrelationships between working conditions, job satisfaction, burnout and mental health among hospital physicians in Japan: A path analysis. Industrial Health. 2008; 47:166–172. [PubMed: 19367046]

Unterbrink T, et al. Burnout and effort-reward imbalance in a sample of 949 German teachers. International Archives of Occupational and Environmental Health. 2007; 80(5):433–441. [PubMed: 17294239]

Van Bogaert P, Clarke S, Roelant E, Meulemans H, Van de Hyning P. Impacts of unit level nurse practice environment and burnout on nurse-reported outcomes: A multilevel modeling approach. Journal of Clinical Nursing. 2010; 19:1664–1674. [PubMed: 20579204]

Waldman SV, Diez JCL, Arazi HC, Linetzky B, Guinjoan S, Grancelli H. Burnout, perceived stress, and depression among cardiology residents in Argentina. Academic Psychiatry. 2009; 33(4):296–301. [PubMed: 19690109]

Wu S, Zhu W, Wang Z, Wang M, Lan Y. Relationship between burnout and occupational stress among nurses in China. Journal of Advanced Nursing. 2007; 59(3):233–239. [PubMed: 17590211]

Potentially Problematic Items in the MBI-HSS from the RN4CAST Cross-Cultural Adaptation Process*^
I feel I treat some patients as if they were impersonal objects.
Working with people all day is really a strain for me.
I worry that this job is hardening me emotionally.
I feel like I'm at the end of my rope.
I feel patients blame me for some of their problems.

* Permission to reprint questions pending.

^ 7 items had low scores but the copyright agreement only permits 5 to be printed.

Fig. 1.
Potentially problematic items, with average modified *k* of 0.60 or lower.

Table 1

Definitions of Flaherty's criteria for evaluating cross-cultural equivalence of survey instrument items.

Criteria	Definition
Content equivalence	The content of each item of the instrument is relevant to the phenomena of each culture being studied.
Semantic equivalence	The meaning of each item is the same in each culture after translation into the language and idiom (written or oral) of each culture.
Technical equivalence	The method of assessment is comparable in each culture with respect to the data that it yields.
Criterion equivalence	The interpretation of the measurement of the variable remains the same when compared with the norm for each culture studied.
Conceptual equivalence	The instrument is measuring the same theoretical construct in each culture.

Adapted from Flaherty et al. (1988), p. 258.

Table 2

Integrative review of studies that translated the MBI-HSS, MBI-GS, & MBI-ES.[a]

Authors	Year	Country	Language	PT	Translation Method	SAM	CE1	SE	TE	CE2	CE3
Ahola, et al.	2006	Finland	Finnish	3+4	MBI-GS – no mention	Gender/age adjusted logistical regression	Y	N	Y	N	?
Al-Dubai and Rumpal	2010	Yemen	Arabic	2	– MBI-HSS "translated into Arabic by a professional translator. The translated version was compared with the English version by the principal author to ensure it reflected original method" – Exploratory and confirmatory factor analysis to confirm accuracy	Descriptive statistics, multiple logistic regression	Y	Y	Y	N	?
Asai et al.	2007	Japan	Japanese	2	No mention, however, "psychometric properties of the Japanese version of the MBI are controversial"	Logical regression analysis	Y	N	Y	N	?
Berg et al.	2006	Norway	Norwegian	4	No mention of translation	Logical regression analysis	Y	N	Y	N	?
Bressi et al.	2009	Italy	Italian	2	States use of "Italian Version"	Linear regression analysis	Y	N	Y	N	?
Chen et al.	2013	Malaysia	Malay	3	MBI-HSS – forward/backward translation. Multi-disciplinary team to reconcile the instruments.	Cronbach's alpha	Y	N	Y	N	?
Córdoba et al.	2011	Columbia	Spanish	1,2 and 3	MBI-HSS – forward/backward translation. Judges evaluated the level of pertinence for each item	Descriptive statistics and Cronbach's alpha	Y	N	Y	N	?
Embriaco et al.	2006	France	French	2	No mention of translation	Ordinal logistical regression	Y	N	Y	N	?
Embriaco et al.	2007	France	French	3	No mention of translation	No mention	Y	N	Y	N	?
Glasberg et al.	2007	Sweden	Swedish	3+4	Validated Swedish translation	Regression analysis	Y	?	Y	N	?
Goehring et al.	2005	Switzerland	French, German, Italian	2	MBI-HSS – validated German and French versions. They created their own version of the Italian.	Logistic regression	Y	Y	Y	N	?
Hu and Schaufeli	2009	China	Chinese	4	MBI-ES – "translated from English into Chinese by three native Chinese speaking master's students…semantic	Confirmatory factor analysis	Y	?	Y	N	?

Authors	Year	Country	Language	PT	Translation Method	SAM	CE1	SE	TE	CE2	CE3
Iglesias et al.	2009	Spain	Spanish	1	differences were agreed upon…" "Both questionnaires have been validated internationally and have been adapted for the Spanish population"	Descriptive statistics	Y	?	Y	N	?
Juthberg et al.	2010	Sweden	Swedish	1	A Swedish translation, but not stated if it is validated.	Partial least squares regression	Y	?	Y	N	?
Kanste et al.	2006	Finland	Finnish	1	MBI-HS – "translation-back translation procedure"	Exploratory and confirmatory factor analysis	Y	?	Y	N	?
Klersy et al.	2007	Italy	Italian	1+2	"Burnout was assessed with the validated Italian-language version of the Maslach Burnout Inventory"	Population-averaged regression models	Y	?	Y	N	?
Lee et al.	2012	Taiwan	Chinese	1	MBI-HSS-forward/backward translation. Pilot study of the Chinese version.	Exploratory and confirmatory factor analysis	Y	Y	Y	N	?
Liakopoulou et al.	2007	Greece	Greek	3	No mention of translation	Descriptive statistics	Y	N	Y	N	?
Luk et al.	2009	Macau	Chinese	4	C-MBI – stated used a Chinese version, but no mention of validity	ANOVA	Y	?	Y	N	?
Mészáros et al.	2013	Hungary	Hungarian	1, 2, and 3	MBI-HSS – forward/backward translation. Pilot test	Confirmatory factor analysis	Y	Y	Y	N	?
Ndetei et al.	2008	Kenya	Swahili	3	MBI HS and GS - no mention of translation	SPSS	Y	?	Y	N	?
Pisanti et al.	2012	Italy	Italian	1	MBI-HSS - forward/backward translation. States it is "substantially equivalent" to another Italian version	Confirmatory factor analysis	Y	N	Y	N	?
Schaufeli et al.	2009	The Netherlands	Dutch	2	They use the Dutch version of the MBI-HSS	Structural equation modeling	Y	Y	Y	N	?
Soler et al.	2006	Many European countries	Many	2	"In those countries where the use of an English-language instrument could potentially pose language barriers, the questionnaire was translated into the native language by a FD"	Descriptive statistics	Y	N	Y	N	?
Tokuda et al.	2009	Japan	Japanese	2	Reliable and valid Japanese version	Path analysis	Y	Y	Y	N	?

Authors	Year	Country	Language	PT	Translation Method	SAM	CE1	SE	TE	CE2	CE3
Unterbrink et al.	2007	Germany	German	4	Use of the MBI-D which is the German version the MBI scale?	Descriptive statistics	Y	N	Y	N	?
Van Bogaert et al.	2009	Belgium	Dutch	1	A previously used translated version of the MBI-HSS, not mention of validity	Structural equation modeling	Y	?	Y	N	?
van der Ploeg et al.	2003	The Netherlands	Dutch	4	They use the Dutch version of MBI	Multiple hierarchical regression analysis	Y	Y	Y	N	?
Waldman et al.	2009	Argentina	Spanish	2	"Employed the Spanish version, which has shown to have adequate reliability and validity in previous studies"	Multivariate logistic regression	Y	?	Y	N	?
Wu et al.	2007	China	Mandarin Chinese	1	MBI-GS-translated from English to Chinese – then back from Chinese to English. The Chinese version was reviewed and was deemed to have high validity	Parametric statistics	Y	?	Y	N	?

[a]*Note:* PT (provider type): RN = 1; MD = 2; Multiple = 3; Other=4; SAM (statistical analysis method); CE1 (content equivalence); CE2 (criterion equivalence); TE (technical equivalence); SE (semantic equivalence); CE1 (content equivalence); CE3 (conceptual equivalence); CE3 (conceptual equivalence); Y = Present; N = Not present; ? = Unable to determine.

Table 3

Expert rater demographics^a ($n = 106$).

	% of Respondents
Education	
Bachelors	63
Masters	37
Role	
Practitioner	50
Educator	29
Administrator	21

^{a}Gender identity of raters was not formally collected for this exercise. Almost all of the raters were female, reflective of the gender dominance within the nursing profession globally.

Table 4

Scale level modified kappa scores by country for the MBI-HSS.

Country	k Score
Belgium (Dutch)	0.84
Belgium (French)	0.49
Finland	0.84
Germany	0.91
Greece	0.93
Ireland	0.51
Netherlands	0.77
Poland	0.6
Sweden	0.81
Switzerland (French)	0.55
Switzerland (German)	0.68
Switzerland (Italian)	0.88
United Kingdom	0.57
United States	0.49
Study Average	0.70

Table 5

Potentially problematic items – a comparison between similar languages.

Country (language)	Belgium (Dutch)		Netherlands (Dutch)		Belgium (French)		Switzerland (French)		Germany (German)		Switzerland (German)		Ireland (English)		UK (English)		US (English)	
Problem item Question	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
..........impersonal objects...............		✓	✓		✓		✓			✓	✓		✓		✓		✓	
..............a strain...............		✓	✓		✓		✓			✓	✓		✓		✓		✓	
..............hardening me..............	✓		✓		✓		✓			✓	✓		✓			✓	✓	
...............don't really care..............		✓		✓	✓		✓			✓	✓		✓		✓		✓	
...............stress..............		✓		✓	✓		✓			✓	✓		✓		✓		✓	
...............end of my rope..............		✓		✓	✓		✓			✓	✓		✓		✓		✓	
...............blame me..............	✓		✓		✓		✓			✓	✓		✓		✓		✓	