Chapter 106

Considerations on the cross-cultural adaptation process and validation of measurement questionnaire

https://doi.org/10.56238/devopinterscie-106

Cláudia Rucco Penteado Detregiachi
Bruna Rezek Andery Altran

1 INTRODUCTION

TRANSLATION AND CROSS-CULTURAL ADAPTATION

Translating an existing instrument demands a shorter period than developing a new one (HUNT et al., 1991). The intention when translating a questionnaire is to develop another version of the instrument with equivalence to the original (HILTON; SKRUTOWSKI, 2002).

Cross-cultural adaptation is defined as the adjustment of a questionnaire to another specific language/dialect of a country or region (AARONSON et al., 2002) and to its respective cultural context and lifestyle (GUILLEMIN; BOMBARDIER; BEATON, 1993).

The translation of instruments from one language to another should not be done in a linear process, since there are several options for translating a word or phrase and there may not be an exactly equivalent translation for a given term (COSTER; MANCINI, 2015). Thus, in this translation process, the instruments should not be literally translated, but culturally adapted to the target population, in order to maintain the validity of their content at a conceptual level even in different cultures (BEATON et al., 2000 ; REICHENHEIM; MORAES, 2007).

For this process to be successfully developed, the translation and cross-cultural adaptation must follow some steps: initial translation; synthesis of translations; back translation; instrument review and pre-test (GUILLEMIN; 1995; BEATON et al., 2000).

The initial translation must be carried out by at least two people working independently, with fluency in the two languages involved, but having the language into which the instrument will be translated as their mother tongue (BEATON et al., 2000; WILD et al., 2005; COSTER; MANCINI, 2015).

After completing the two independent translations, they must proceed to their synthesis stage in a reconciliation process that involves comparing the two versions to identify points of difference and then reconciling them to create a single version (WILD et al., 2005 ; COSTER; MANCINI, 2015). This process can be developed in different ways. One possible approach is to form a committee of experts (BEATON et

al., 2000) to review and discuss discrepancies until reaching a consensus on each point (BEATON et al., 2000; COSTER; MANCINI, 2015). These specialists must analyze item by item, choosing the best translation and, when necessary, suggest adaptations, always focusing on cultural and linguistic characteristics that may cause difficulties when the original language version was translated into the new language (BRACCIALLI et al., 2016). It is important for the research coordinator to participate in reconciliation decisions in order to allow a certain degree of consistency and harmonization with other translated versions (GUILLEMIN, 1995).

From the conciliation process, a consensus is reached on the final version of the translated instrument, which must then be submitted to the back-translation or reverse-translation stage. This step aims to assess whether the translated version reflects the original version (BORSA; DAMÁSIO; BANDEIRA, 2012), when then the translated version is translated into its original language (COSTER; MANCINI, 2015). This reverse translation must be performed by a person who is fluent in both languages, but with the native language of the instrument's original language and without knowledge of its existence (WILD et al., 2005). This translator should not know the objectives and concepts underlying the study (MELCHIORS; CORRER; FERNÁNDEZ-LLIMOS, 2007). There is no consensus on the number of back-translations that should be performed, since some guidelines suggest more than one back-translation, to be performed in parallel or sequentially, others suggest a back-translation panel and many propose a single back-translation (WILD et al., 2005) .

In the last step of the translation process, the back-translated version must be sent to the author and/or original editor of the instrument for a review, with the aim of ensuring that the translation is accurate and that it has maintained the most important characteristics of the original instrument (COSTER; MANCINI, 2015). This review seeks to verify the semantic, idiomatic and conceptual equivalence between the original instrument and the back-translated one from the perspective of the referential meaning of the constituent terms and words (similarity regarding the literal meaning of the constituent terms of the pairs of assertions) and regarding the general meaning of each item (similarity regarding the idea conveyed by the assertions) (HERDMAN; FOX-RUSHBY; BADIA, 1998; SPERBER, 2004). Semantic equivalence is evaluated based on grammar and vocabulary, while idiomatic equivalence is based on colloquialism or difficult-to-translate expressions (BEATON et al., 2000). In this step, the author and/or editor may find terms in the reverse translation that are not identical to those in the original version. In this case, he can discuss with the research team and help to understand the meaning of the original term and then define the most equivalent term in the translation language (COSTER; MANCINI, 2015). At the end of this step, the translated version of the instrument is available.

In the final stage of the cross-cultural adaptation process, the instrument in the translated language version must be submitted to a pre-test that consists of applying the instrument to about 12 to 15 individuals who have characteristics similar to the target sample (BOTELHO, 2007). Here, the objective is to verify the clarity of the instrument in relation to the understanding of the items that the

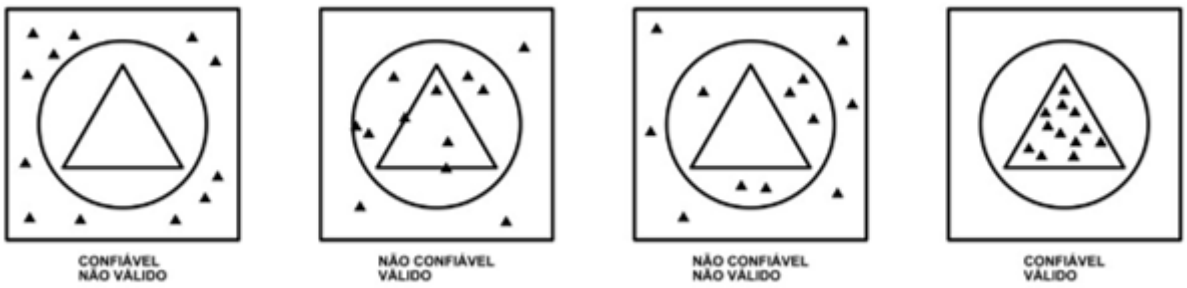
make up. The instrument is understood by at least 90% of the participants. Otherwise, the translated version must undergo adjustments by the researchers and a new consultation with the expert committee must be carried out, with the version in question being reformulated and submitted to a new pre-test until such an index of comprehension is reached (DE SOÁREZ et al ., 2007).

2 ANALYSIS OF PSYCHOMETRIC PROPERTIES

After achieving equivalence between the original source and the translated version, through translation and cross-cultural adaptation, it is necessary to study its validity so that the instrument can be used in a new culture (BEATON et al., 2000), which is developed based on the principles of psychometrics.

The psychometric properties verify the quality of the information provided by the instrument (ROACH, 2006), with reliability and validity being the main measurement properties of instruments (PITTMAN; BAKAS, 2010; CANO; HOBART, 2011; SOUZA; ALEXANDRE; GUIRARDELLO, 2017). Reliability is the ability to reproduce a result consistently, in time and space; and validity refers to the property of an instrument to measure exactly what it proposes (SOUZA; ALEXANDRE; GUIRARDELLO, 2017) (Figure 1).

Figure 1 - Possible combinations of validity and reliability properties of measurement instruments. Source: the author adapted from Babbie (1986).



The guidelines for the sample size of participants in the validation process ranged from 5:1 (GORUSCH, 1983), that is, five respondents for each item of the questionnaire; 10:1 (NUNNNALLY, 1978); 15:1 and even 30:1 (PEDHAZUR, 1997). There is also a recommendation that a sample of 50 participants be considered as very poor, 100 as poor, 200 as fair, 300 as good, 500 as very good and 1000 or more as excellent (COMFREY; LEE, 1992).

3 RELIABILITY

Reliability, or trustworthiness, is the ability of the instrument to consistently reproduce a result in time and space, or from different observers, indicating aspects of coherence, precision, stability, equivalence and homogeneity. This is one of the main quality criteria of an instrument (TERWEE et al., 2007), which is evaluated using two criteria, stability and internal consistency.

Stability

The stability of a measure is the degree to which similar results are obtained at two different times (POLIT; BECK, 2011), that is, it is the estimate of the consistency of the repetitions of the measures (SOUZA; ALEXANDRE; GUIRARDELLO, 2017).

This measurement is performed using the test-retest method, which consists of applying the instrument under study twice by different interviewers (E1 and E2), in the same period, but with an interval of 30 to 60 minutes (inter-observer measurement). This application is repeated by one of the interviewers (E2) in an interval of 10 to 14 days (intra-observer measurement) (KESZEI; NOVAK; STREINER, 2010), which is enough time to exclude the memory effect, but not to generate change in the measured factor (CARDOSO, 2007; VIEHWEGER et al., 2008; POLIT; BECK, 2011; ECHEVARRÍA-GUANILO; GONÇALVES; ROMANOSKI, 2017).

The intraclass correlation coefficient (ICC) is the statistical test used to estimate the stability of continuous variables (VET et al., 2006; TERWEE et al., 2011), with values greater than or equal to 0.75 are considered excellent, satisfactory between 0.4 and 0.75 and poor when less than 0.4 (HULLEY et al., 2001). The use of this method requires that the factor to be measured remains the same in the two moments of the tests and any change in the score can be caused by random errors (KESZEI; NOVAK; STREINER, 2010).

Internal Consistency

Internal consistency, or homogeneity, indicates whether all subparts of an instrument measure the same characteristic (NUNNALLY, 1994). This is an important measurement property for instruments that assess a single construct, using a variety of items (TERWEE et al., 2007). A low internal consistency estimate may mean that the items measure different constructs or that the answers to the instrument's questions are inconsistent (KESZEI; NOVAK; STREINER, 2010).

To assess the internal consistency of the instrument under study, Cronbach's alpha coefficient is used, which reflects the degree of covariance between its items. Thus, the smaller the sum of the variance of the items, the more consistent the instrument is considered (PASQUALI, 2013). For interpretation, values above 0.70 are considered ideal (NUNNALLY; BERNSTEIN, 1994; TERWEE et al., 2007), although Maroco, Garcia-Marques (2006) recommend values above 0.5 for measurements with reliability to compare groups of individuals.

4 VALIDITY

Validity refers to the fact that an instrument measures exactly what it is intended to measure (ROBERTS; PRIEST, 2006; MOKKINK et al., 2010). This is evaluated under three aspects: content, criterion and discriminant validity.

Content Validity

Content validity refers to the degree to which the content of an instrument adequately reflects the construct being measured (POLIT, 2015), that is, it is the assessment of how much a sample of items is representative of a defined universe or domain. of a content (POLIT, BECK, 2011). For example, an instrument that assesses job satisfaction should include not only satisfaction but also other variables related to it, for example, remuneration, promotion, relationships with co-workers, among others (SOUZA; ALEXANDRE; GUIRARDELLO, 2017) . Reaffirming, Ramada-Rodilla; Serra-Pujadas; Delclós-Clanchet (2013) define content validity as the degree to which the content of an instrument is able to measure most of the dimensions of the evaluated construct. These researchers add that a questionnaire with high content validity is one that measures all dimensions related to the theoretical construction in question.

Given the fact that there is no specific statistical test to measure content validity, a qualitative approach is used, through the evaluation of a committee of experts (KIMBERLIN; WINTERSTEIN, 2008) and after a quantitative approach using the index of content validity (CVI) (COLUCI; ALEXANDRE; MILANI, 2015), which measures the proportion or percentage of judges in agreement on certain aspects of the instrument and its items (ALEXANDRE; COLUCI, 2011).

Criterion validity

Criterion validity consists of the relationship between scores of a given instrument and some external criterion (KIMBERLIN; WINTERSTEIN, 2008). This criterion should consist of a widely accepted measure, with the same characteristics as the assessment instrument, that is, an instrument or criterion considered “gold standard” (KESZEI; NOVAK; STREINER, 2010).

In criterion validity assessments, the validity of a measure is tested by comparing the measurement results with a “gold standard” or established criterion. When the target test measures what it is intended to measure, its results must agree with the “gold standard” or criterion results (ROACH, 2006). Whatever the evaluated construct, the new instrument is considered valid when its scores match the scores of the chosen criterion (POLIT, BECK, 2011), that is, the “gold standard”.

Criterion validity is verified by the correlation coefficient (POLIT, BECK, 2011), with the measurement instrument scores correlated with the “gold standard” scores (KESZEI; NOVAK; STREINER, 2010). It is considered that values close to 1.00 indicate that there is a correlation, while values close to 0.00 indicate its absence, with correlation coefficients of 0.70 or higher being desirable (POLIT, BECK, 2011).

Discriminant validity

Discriminant validity tests the hypothesis that the measure in question is not unduly related to different constructs, that is, to variables from which it should differ. This measure is intended to verify, for example, whether the scores of the instrument being studied for validation differ between patients with

“bad and good” metabolic controls (POLIT, 2015), or even whether an instrument that assesses motivation for work presents low correlations with an instrument that verifies self-efficacy at work (AARONSON et al., 2002). To evaluate this aspect, the statistical correlation test is used (POLIT, 2015).

REFERENCES

AARONSON, N.; ALONSO, J.; BURNAM, A. et al. Assessing health status and quality-of-life instruments: attributes and review criteria. ***Qual Life Res***., v. 11, n. 3, p. 193-205, 2002.

ALEXANDRE, N.M.C.; COLUCI, M.Z.O. Validade de conteúdo nos processos de construção e adaptação de instrumentos de medidas. **Cienc Saude Coletiva**, v. 16, n. 7, p. 3061-3068, 2011.

BABBIE, E. **The practice of social research**. 4th ed. Belmont: Wadsworth Publishing Company; 1986.
BEATON, D.E. et al. Guidelines for the Process of Cross Cultural Adaptation of Self-Report Measures. **Spine**, v. 25, n. 24, p. 3186-91, 2000.

BORSA, J.C.; DAMÁSIO, B.F.; BANDEIRA, D.R. Adaptação e Validação de Instrumentos Psicológicos entre Culturas: Algumas Considerações. **Paidéia**, v. 22, n. 53, p. 423–432, 2012.

BOTELHO A.C.S.R. **Validação para o uso no brasil do "Pediatric Quality of Life™" (PEDSQLTM):** um estudo envolvendo famílias com crianças e adolescentes acometidos por câncer. 2007. Dissertação (Mestrado em Odontologia). Faculdade de Odontologia, Universidade Federal de Minas Gerais, Belo Horizonte, 2007.

BRACCIALLI, L.M. et al. Translation and validation ofthe Brazilian version of the Cerebral Palsy Quality of Life Questionnaire for Children - child report. **J Pediatr**, v. 92, n. 2, p. 143-8, 2016.

CANO, S.J.; HOBART, J.C. The problem with health measurement**. Patient Prefer Adherence**. v. 5, p. 279-90, 2011.

CARDOSO, M.A. Desenvolvimento, Validação e Aplicações de Questionários de Frequência Alimentar em Estudos Epidemiológicos. *In*: KAC, G.; SICHIERI, R.; GIGANTE, D.P. (orgs.) **Epidemiologia nutricional** [online], Rio de Janeiro: Editora FIOCRUZ/Atheneu, 2007. p. 200-12.

COLUCI, M.Z.O.; ALEXANDRE, N.M.C.; MILANI, D. Construção de instrumentos de medida na área da saúde. **Cienc Saude Coletiva**, v. 20, n. 3, p. 925-36, 2015.

COMFREY, A.L.; LEE, H.B. **A First Course in Factor Analysis. Hillsdale**, NJ: Lawrence Erlbaum Associates; 1992.

COSTER, W.; MANCINI, M. Recomendações para a tradução e adaptação transcultural de instrumentos para a pesquisa e a prática em Terapia Ocupacional. **Rev Ter Ocup Univ São Paulo**, v. 26, n. a, p. 50-57, 2015.

DE SOÁREZ, P.C. et al. Tradução para português brasileiro e validação de um questionário de avaliação de produtividade. **Rev Panam Salud Publica**. v. 22, n. 1, p. 21-8, 2007.

ECHEVARRÍA-GUANILO; M.E.; NATÁLIA GONÇALVES, N.; ROMANOSKI, P.J. Propriedades psicométricas de instrumentos de medidas: bases conceituais e métodos de avaliação – Parte I. **Texto Contexto Enferm**, v. 26, n. 4, p. 1-11, 2017.

GORUSCH, R.L. **Factor Analysis**. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1983.

GUILLEMIN, F. Cross-cultural adaptation and validation of health status measures. **Scand J Rheumatol.**, v. 24, n. 2, p. 61-63, 1995.

GUILLEMIN, F.; BOMBARDIER, C.; BEATON, D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. **J Clin Epidemiol**, v. 46, n. 12, p. 1417-32, 1993.

HERDMAN, M.; FOX-RUSHBY, J.; BADIA, X. A model of equivalence in the cultural adaptation of HRQoL instruments: the universalist approach. **Qual Life Res**, v. 7, p. 323-35, 1998.

HILTON, A.; SKRUTOWSKI, M. Translating instruments into other languages: development and testing processes. **Cancer Nurs.**, v. 25, p. 1-7, 2002.

HULLEY, S.B. et al. **Designing clinical research: an epidemiologic approach**. Philadelphia, PA: Lippincott Williams & Wilkins; 2001. 252 p.

HUNT, S.M., et al. Cross-cultural adaptation of health measures. European Group for Health Management and Quality of Life Assessment. **Health Policy**, v. 19, n. 1, p. 33-44, 1991.

KESZEI, A.P.; NOVAK, M.; STREINER, D.L. Introduction to health measurement scales. **J Psychosom Res**. V.68, n. 4, p. 319-23, 2010.

KIMBERLIN, C.L.; WINTERSTEIN, A.G. Validity and reliability of measurement instruments used in research. **Am J Health Syst Pharm**., v. 65, n. 23, p. 2276-84, 2008.

MAROCO, J.; GARCIA-MARQUES, T. Qual a fiabilidade do alfa de Cronbach? Questões antigas e soluções modernas. **Laboratório de Psicologia**, v.4, n. 1, p. 65-90, 2006.

MELCHIORS, A. C.; CORRER, C. J.; FERNÁNDEZ-LLIMOS, F. Tradução e Validação para o Português do Medication Regimen Complexity Index. **Arquivos Brasileiros de Cardiologia**, v. 89, n. 4, p. 210–218, 2007.

MOKKINK, L.B. et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. **J Clin Epidemiol**., v. 63, n. 7, p. 737-75, 2010.

NUNNALLY, J. **Psychometric Theory**. New York: McGraw-Hill; 1978.

NUNNALLY, J.C.; BERNSTEIN, I.H. **Psychometric theory**. 3rd. Ed. New York: McGraw-Hill; 1994.

PASQUALI, L. **Psicometria**: teoria dos testes na psicologia e na educação. 1 ed. Rio de Janeiro: Vozes; 2013.

PEDHAZUR, R.J. **Multiple Regression in Behavioral Research**: Explanation and Prediction. Fort Worth, TX: Harcourt Brace College Publishers; 1997.
PITTMAN, J.; BAKAS, T. Measurement and instrument design. **J Wound Ostomy Continence Nurs**., v. 37, n. 6, p. 603-7, 2010.

POLIT, D.F. Assessing measurement in health: beyond reliability and validity. **Int J Nurs Stud**., v. 52, n. 11, p. 1746-53, 2015.

POLIT, D.F.; BECK, C.T. **Fundamentos de pesquisa em enfermagem**: métodos, avaliação e utilização. 7ª ed. Porto Alegre: Artmed; 2011.

RAMADA-RODILLA, J.M.; SERRA-PUJADAS, C.; DELCLÓS-CLANCHET, G.L. Adaptación cultural y validación de cuestionarios de salud: revisión y recomendaciones metodológicas. **Salud Publica Mex**, v. 55, n. 1 p. 57-66, 2013.

REICHENHEIM M. E.; MORAES C. L. Operacionalização de adaptação transcultural de instrumentos de aferição usados em epidemiologia. **Rev Saúde Pública**, v. 41, n. 4, p. 665-673, 2007.

ROACH, K.E. Measurement of health outcomes: reliability, validity and responsiveness. **J Prosthet Orthot**., v. 18, n. 6, p. 8-12, 2006.

ROBERTS, P.; PRIEST, H. Reliability and validity in research. **Nurs Stand**., v. 20, n. 44, p. 41-5, 2006.

SOUZA, A.C.; ALEXANDRE, N.M.C.; GUIRARDELLO, E.B. Propriedades psicométricas na avaliação de instrumentos: avaliação da confiabilidade e da validade. **Epidemiol. Serv. Saude**, v. 26, n. 3, p. 649-59, 2017.

SPERBER, A.D. Translation and Validation of Study Instruments for Cross-Cultural Research. **Gastroenterology**, v. 126, p. S124-S128, 2004. Suppl. 1.

TERWEE, C.B. et al. Methodological quality of studies on the measurement properties of neck pain and disability questionnaires: a systematic review. **J Manipulative Physiol Ther**. v. 34, n. 4, p. 261-72. 2011.

TERWEE; C.B. et al. Quality criteria were proposed for measurement properties of health status questionnaires. **J Clin Epidemiol**., v. 60, n. 1, p. 34-42, 2007.

VET, H.C. et al. When to use agreement versus reliability measures. **J Clin Epidemiol**., v. 59, n. 10, p. 1033-39, 2006.

VIEHWEGER, E. et al. Measuring quality of life in cerebral palsy children. **Ann Readapt Med Phys**., v. 51, n. 2, p. 129-37, 2008.

WILD, D. et al. Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. **Value Health**., v. 8, n. 2, p. 94-104, 2005.