Editorial

International Journal of Psychiatry

Cross Cultural Adaptation & Psychometric Validation of Instruments: Step-wise Description

Arafat S M Yasir[*12]

[1]*Department of Psychiatry, Bangabandhu Sheikh Mujib Medical University, Dhaka, Bangladesh*

[2]*Department of Public Health, ASA University Bangladesh*

[*]**Corresponding author**
Arafat S M Yasir, Department of Psychiatry, Bangabandhu Sheikh Mujib Medical University, Dhaka, Bangladesh, Tel: +8801713272917; E-mail: arafatdmc62@gmail.com.

Submitted: 04 July 2016; **Accepted**: 06 July 2016; **Published**: 09 July 2016

Abstract

Psychometrics has immense role in measuring health outcomes across the globe. To use the scales in different culture, there is need to adapt the instrument appropriately. It was aimed to describe the psychometric validation of instruments step by step as well as comprehensively. Cultural adaptation needs to follow standard adaptation procedure and assessment of reliability and validity after ensuring the standard sample size. Reliability is assessed mostly in the forms of internal consistency; test-retest reliability; and inter rater reliability. Validity is assessed in forms of face validity; content validity; criterion validity; construct validity; concurrent validity; convergent validity and divergent validity. Study design; sample size estimation and sampling technique vary depending on the situation and demands thorough scientific discussions. This step wise comprehensive description will make a comfortable path for the beginners.

Keywords: Scale validation, Validation methodology, Psychometric validation, Cross-cultural Adaptation.

Introduction

Psychometrics has immense role in psychiatry, public health, primary health care, and many others fields; even in health promotional strategy for measuring the attitude [1]. To use any scales in different culture rather than its origin, there is need to adapt the measuring instrument appropriately. It is important to realize that self reporting scales are potentially vulnerable to distortion due to a range of factors, including social desirability, dissimulation, and response style [2]. Consequently, there is much emphasis on using standardized and validated research instruments to measure the responses [3]. Moreover, culturally adapting of an instrument has many advantages over developing a new one such as, it reduces the costs and the time spent in development [4]. Cultural adaptation and psychometric validation comprised of series of process stating with cultural adaption by following standard procedure; followed by assessing the different forms of reliability and validity by well accepted scientific measures. However, to author's best knowledge, there is no comprehensive guideline to follow during cultural adaptation and psychometric validation of measuring scales as well as there is paucity of literatures describing the every steps for the beginners. So, author aimed to describe the process step wise as well as comprehensively so that the beginners have a comfortable journey in the validation process.

Cross cultural adaptation

Cross-cultural adaptation of an instrument for use in a new country, culture, and/or language necessitates use of a unique method, to reach equivalence between the original source and target versions of the instrument [5]. Guideline described by Beaton, et al. is the mostly used and practiced guideline; having steps of initial translation by minimum two translators who have adequate understanding regarding the both languages, among them one will be informed regarding the process and other will be disguised; synthesis of the translations by resolving the differences between the translations, better to compiled by third another person; back translation of the compiled translation by at least two persons who have good understanding regarding the both languages; expert committee comprised of methodologists, health professionals, language professionals, and the translators (forward and back-translators) will review the steps and will consider the semantic, idiomatic, experiential and conceptual equivalences; pretesting of the reviewed questionnaire have to done at least 30-40 individuals and suggested changes should be considered; & finally questionnaire is accepted for collecting the responses (Figure 1) [3-5]. Translation is the first stage of the adaptation process but the term "adaptation" means different from "translation" as adaptation includes all the processes concerning cultural, idiomatic, linguistic and contextual aspects concerning its translation [6,7]. Adaptation ensures demanded equivalences covering semantic, idiomatic, experiential and conceptual equivalence backed by the expert committee review. Semantic Equivalence; ensures the

equivalence of meaning as the translated version needs to mean the same with the original. Idiomatic Equivalence; ensures the equivalence of colloquialisms, or idioms, are difficult to translate. Experiential Equivalence; ensures the experiential quality of the translated questionnaire in regards to the items aiming to capture and experience of daily life which may have differences from the original version. Conceptual Equivalence; ensures the conceptual meaning replacement that are different from culture to culture. The steps mentioned in Figure 1 are the best guided adaptation process fetching the consideration of standard translation, adequate equivalency and furthermore changes supported by the pretesting [3-5,8].

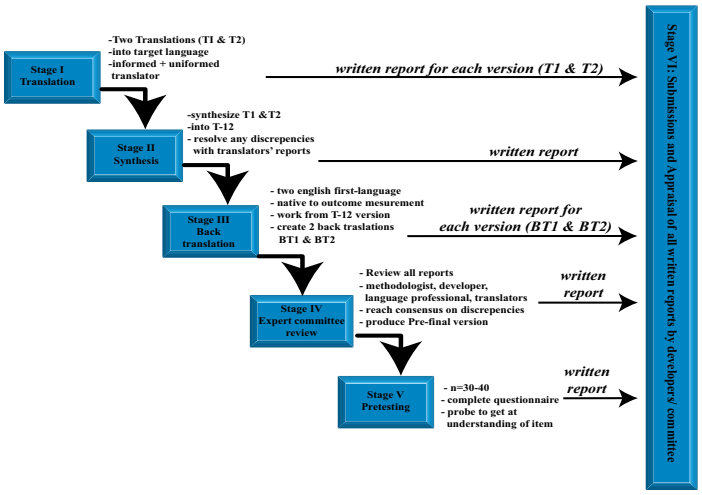

Figure 1: Graphical representation of translation process for validation study [5].

Reliability assessment

Reliability of a measure refers to the ability of a questionnaire to determine that a measurement yields reproducible and consistent results [4,8]. Reliability can be analyzed mostly in the forms of internal consistency, test-retest reliability and inter-rater reliability [4,8,9] (Table 1).

Table 1: Definitions of psychometric properties of a scale.

Property	Definition
Face validity	The ability of an instrument to be understandable and relevant for the targeted population [4,9].
Content validity	The ability of an instrument to reflect the domain of interest and the conceptual definition of a construct [4,9].
Construct validity	The extent to which a measure is related to specified variables in accordance with an established theory or hypothetical construct' [4,9].
Convergent validity	The degree to which scores on a measure associate with scores on other measures that are intended to assess similar constructs [4,9].
Divergent validity	Involves that items within any one subscale should not correlate too highly with external items or with the total sum-score of another subscale [4,9].
Discriminative validity	The ability of an instrument to distinguish between groups that are expected to differ based on their clinical diagnosis or other characteristics [4].
Criterion validity	The assessment of an instrument against the true value, or a standard accepted as the true value [4,9].
Concurrent validity	The association of an instrument with accepted standards [4,9].
Internal consistency	The ability of an instrument to have interrelated items [4,9].
Test-retest reliability	The ability of the scores of an instrument to be reproducible if it is used on the same patient while the patient's condition has not changed (measurements repeated over time) [4,9].

Internal consistency is assesses by Cronbach's alpha having level of ≥0.70 is considered as significant [1,4,8] and the measure is preloaded in the analyzing software such as in SPSS in the data reduction menu under analyzing toolbar. Test-retest reliability is assessed by applying same instrument to the same respondents after a certain time period; usually after the two weeks [4,8]. It can be assessed by different statistical measures such as Wilcoxon Non-parametric Statistical Test, intra-class correlation coefficient (ICC), Pearson's correlation coefficient, Spearman Correlation Coefficient, Kappa coefficients, paired t tests; that structured in the software; based on the situation as well as researcher's criteria [4]. Inter-rater reliability is assessed by applying the same instrument by two or more raters and comparing the responses; can be done by Kappa statistics, Student's t test statistical analysis and other measures on basis of the available situation and researchers' criteria; those are preloaded in commercial software [4].

Validity assessment

Validity of a measure refers to the ability of a instrument to measure what it is supposed to measure [4]. Validity can be assessed in different forms such as face validity, content validity, construct validity, criterion validity, convergent validity, divergent validity, concurrent validity and others [4,8,9] (Table 1). The validity assessment may be customized on the basis of researchers' criteria, instrument feasibility and statistical analysis. Face validity; is the easiest and weakest form of validity and it can be checked by ensuring standard back translation process during the translation and adaptation process; by critical review and expert panel opinion, feasibility, readability, consistency of style and formatting, and the clarity of the language used [4,8-10]. Content validity; is measured as well as ensured by ensuring standard back translation process; by literature review and expert panel opinion; and by experts with content validity index [4,8,10,11]. Construct validity; can be assessed on basis of the statistical analysis done by factor analysis available in the software; and by comparing with other similar instruments [1,4,8-10,12]. Criterion validity; can be assessed by comparison of the instrument with the gold standard instrument with few variations [4,8,9]. Convergent validity & Concurrent validity; can be ascertain by comparing with other similar instruments with the help of the statistical measures such as Pearson's correlation coefficient, Spearman Correlation Coefficient, etc [4,8].

Factor analysis

Factor analysis significantly contributes in the study process, which includes exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) with having freedom to choose for the researchers'. Factor analysis can help to ascertain sampling adequacy, internal consistency, factor rotation, factor identification, item retention and other steps of analysis; those are predesigned in the available software. EFA in form of Principal Component Analysis (PCA) with Varimax rotation is by far the most common choice to assess the structure of the construct. EFA allows the researcher to explore the main dimensions to generate a theory, or model from a relatively large set of latent constructs often represented by a set of items; whereas CFA is used to test a proposed theory [1,4,8,13].

Factor retention

A construct may have multiple factors or a construct may be uni dimensional. With the help of the factor analysis in form of Principal Component Analysis (PCA) with Varimax rotation, researchers can ascertain the factors in the construct. Factors having eigenvalue of ≥ 1 are used to retention; whereas the graphical representation of eigenvalues named as Scree plot can also be used to retain the factors; and both of the measures are preloaded in the software [4].

Item reduction

To reduce items from the construct researchers can use the factor analysis in varimax rotation. Items can be discarded having value <0.30, although in some instances these criteria can be relaxed [4,13,14].

Study design

Regarding the study design, separate study design such as "validation study" or "methodological study" can be used instead of cross-sectional study design; as the process of scale validation follows distinct scientific steps in adaptation, inter rater reliability assessment, test retest reliability assessment and sample size estimation; those are different from the cross-sectional study design [4,8].

Sampling technique & sample size

Both probability and non probability sampling techniques can be chosen as sampling technique based on the feasibility and approachable accepted methods. There are few accepted methods of estimating sample size without having fixed guidelines. Recent mostly used option is item sample ratio where majority of the authors use the method for sample size estimation with an arbitrary margin of 2 to 20 and reviews showed that subject to item ratios of ≤ 10:1 covers 63.2% studies [1,4,8,9]. Sample size also can be estimated on basis of Exploratory Factor Analysis (EFA), where there are recommendations to ensure the sample size as mentioned; 100 = poor, 200 = fair, 300 = good, 500 = very good, ≥1000 = excellent [4,8,9,13]. The third used approached based on the regression formula proposed by Walter et al, which is mostly on the reliability approach and the sample size vary based on the researchers criteria but it's difficult to estimate the samples where no inter rater reliability as well as the rest retest reliability cannot be performed [15]. Kaiser-Meyer-Olkin (KMO) sampling adequacy test can be performed as a statistical significance of sampling size as it is preloaded in the analyzing software such as Statistical Package for Social Science (SPSS) in the factor analysis menu. The KMO value ≥ 0.6 can be taken as significant [1,4,8,13].

Conclusion

Cross-cultural validation of health instruments is an important area to address comprehensively. It is aimed to describe the cross-cultural adaptation and psychometric validation in stepwise, in brief and comprehensively that can be helpful for the beginners in psychometrics.

References

1. Arafat SM (2016) Psychometric Validation of the Bangla Version of the Patient-Doctor Relationship Questionnaire. Psychiatry J 2016: 9385364.
2. Lenderink AF, Zoer I, van der Molen HF, Spreeuwers D, Frings-Dresen MH, et al. (2012) Review on the validity of self-report to assess work-related diseases. Int Arch Occup Environ Health 85: 229-251.
3. Gjersing L, Caplehorn JR, Clausen T (2010) Cross-cultural adaptation of research instruments: language, setting, time and statistical considerations. BMC Med Res Methodol 10: 13.
4. Arafat SMY, Chowdhury HR, Qusar MS, Hafez MA (2016) Cross Cultural Adaptation & Psychometric Validation of Research Instruments: a Methodological Review. J Behav Heal 5: 3.
5. Beaton DE, Bombardier C, Guillemin F, Ferraz MB (2000) Guidelines for the process of cross-cultural adaptation of self-report measures. Spine (Phila Pa 1976) 25: 3186-3191.
6. Semage SN, Sivayogan S, Forbes D, O'Donnell M, Monaragala RM, et al. (2013) Cross-cultural and factorial validity of PTSD check list-military version (PCL-M) in Sinhalese language. Eur J Psychotraumatol 4.
7. Borsa JC, Damasio BF, Bandeira DR (2012) Cross-Cultural Adaptation and Validation of Psychological Instruments: Some Considerations. Paideia 22: 423-432.
8. Arafat SMY (2016) Validation study can be a separate study design. Int J Med Sci Public Health 5: 11.
9. Anthoine E, Moret L, et al. (2014) Sample size used to validate a scale: a review of publications on newly-developed patient reported outcomes measures. Health Qual Life Outcomes 12: 176.
10. Parsian N, Dunning Am T (2009) Developing and Validating a Questionnaire to Measure Spirituality: A Psychometric Process. Glob J Health Sci 1.
11. Fu SN, Chin WY, Wong CK, Yeung VT, Yiu MP, et al. (2013) Development and validation of the Chinese Attitudes to Starting Insulin Questionnaire (Ch-ASIQ) for primary care patients with type 2 diabetes. PLoS One 8: e78933.
12. Kaiser U, Steinmetz D, Scharnagel R, Jensen MP, Balck F, et al. (2014) Cross-cultural adaptation, evaluation and validation

of the Spouse Response Inventory: a study protocol. BMJ Open 4: e005119.
13. Costello AB, Osborne JW (2005) Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most from Your Analysis. Pract Assessment, Res Eval 10: 9.
14. Atkins R (2014) Instruments measuring perceived racism/ racial discrimination: review and critique of factor analytic techniques. Int J Health Serv 44: 711-734.
15. Walter SD, Eliasziw M, Donner A (1998) Sample size and optimal designs for reliability studies. Stat Med 17: 101-110.