Chapter

Psychometric Analyses in the Transcultural Adaptation of Psychological Scales

Guillaume Gronier

Abstract

Measurement scales play an important role in the methodology of psychological research and practice. They make it possible to obtain scores linked to numerous individual characteristics (feeling of hope, perceived stress, experience, felt well-being, etc.) and thus to draw up a profile of respondents or to compare several situations with each other according to their psychological impact. Most of the research on the construction of these scales is Anglo-Saxon and, therefore, proposes scales in English. However, many non-English speaking countries feel the need to use these scales for their studies, which requires them to be translated into a target language. This proposed chapter describes the steps and psychometric analyses required to adapt an English scale in another language. Based in particular on the recommendations of the International Test Commission and the APA Standards of Practice for Testing, this chapter aims to guide researchers who wish to undertake the translation of a psychological scale. It also includes an analysis of the literature on the translation practices of some one hundred scales, translated and published recently in various scientific journals.

Keywords: translation, questionnaire, scale, psychometric analyses

1. Introduction

Psychology has long mobilised the subjective assessment of individual characteristics using questionnaires or measurement scales. These self-administered scales, i.e., which subjects are invited to respond alone, capture the perception that subjects have of themselves. Without being exhaustive, this may, for example, concern their perceived well-being or ill-being, their perception of certain personality traits, their satisfaction with a product, or their way of apprehending a particular situation. These scales generally have a diagnostic purpose: they provide a score that, once interpreted, gives an evaluation of the subject's perception. While some scales, particularly in the health field, propose thresholds for interpreting their scores, most of them leave the researcher or practitioner free to interpret the meaning of the scores obtained.

The design of these scales is based on a very specific scientific approach, which generally follows the Churchill paradigm [1]. The methodological paradigm for scale construction defined by Churchill aims not only to reduce the common biases in scale completion (halo bias, social desirability bias, contamination bias and response

polarisation bias) but also to verify the internal validity of the scale. The approach is thus based on a succession of stages of item definition, data collection and psychometric analysis, which, as part of an iterative process, ultimately makes it possible to validate the scale that has been designed. Some psychology scales have been validated and used for many years. For example, the Depression Anxiety Stress Scales (DASS) [2] have been used for over 25 years to measure perceived stress and anxiety in clinical psychology.

Therefore, when research requiring the use of certain psychological scales is conducted in languages other than that of the original scale, it seems simpler and more reliable to translate these scales than to create new ones from scratch in the target language. Adapting a scale into new languages thus has the following advantages:

- to be able to compare the scores of the same scale submitted to different audiences of different languages;
- to draw on scales that have often been used in different research studies, thus allowing the advantages and limitations of the scale to be better delineated;
- to continue psychometric analyses of the original scale in new studies; and
- promote research on scale construction.

Like the creation of a new scale, the cross-cultural adaptation of a scale is based on a clearly defined process, of which there are two main steps: the translation of the scale into the target language and the analysis of the psychometric properties of the translation. From a psychometric point of view, the aim is to ensure that the translated version corresponds to the properties of the original version, with particular attention paid to factor correspondence.

This chapter aims to summarise the psychometric analyses necessary for the validation of cross-cultural adaptations of psychology scales. It is thus intended as an aid to researchers and practitioners who wish to adopt a scale into a new language.

2. General methodology for cross-cultural adaptation of psychology scales

Several methodological frameworks describe the steps necessary for cross-cultural adaptation and validation of scales [1–4]. These frameworks are regularly discussed and adapted to provide a more reliable methodology. One of the most common frameworks is the one proposed by the International Test Commission, called the ITC Guidelines for Translating and Adapting Tests (Second Edition) [5]. This guide provides a set of 18 recommendations for conducting and evaluating the adaptation (sometimes also referred to as 'localisation') or simultaneous development of psychological and educational tests for use with different populations.

The 18 recommendations are divided into six main themes: preconditioning, test development, confirmation, empirical analyses, administration, score scales and interpretation and documentation. **Figure 1** summarises the framework described by the ITC.

Among these steps, some require psychometric treatments for the validation of the scales during cross-cultural adaptation, in particular, step 5 'Score scales and

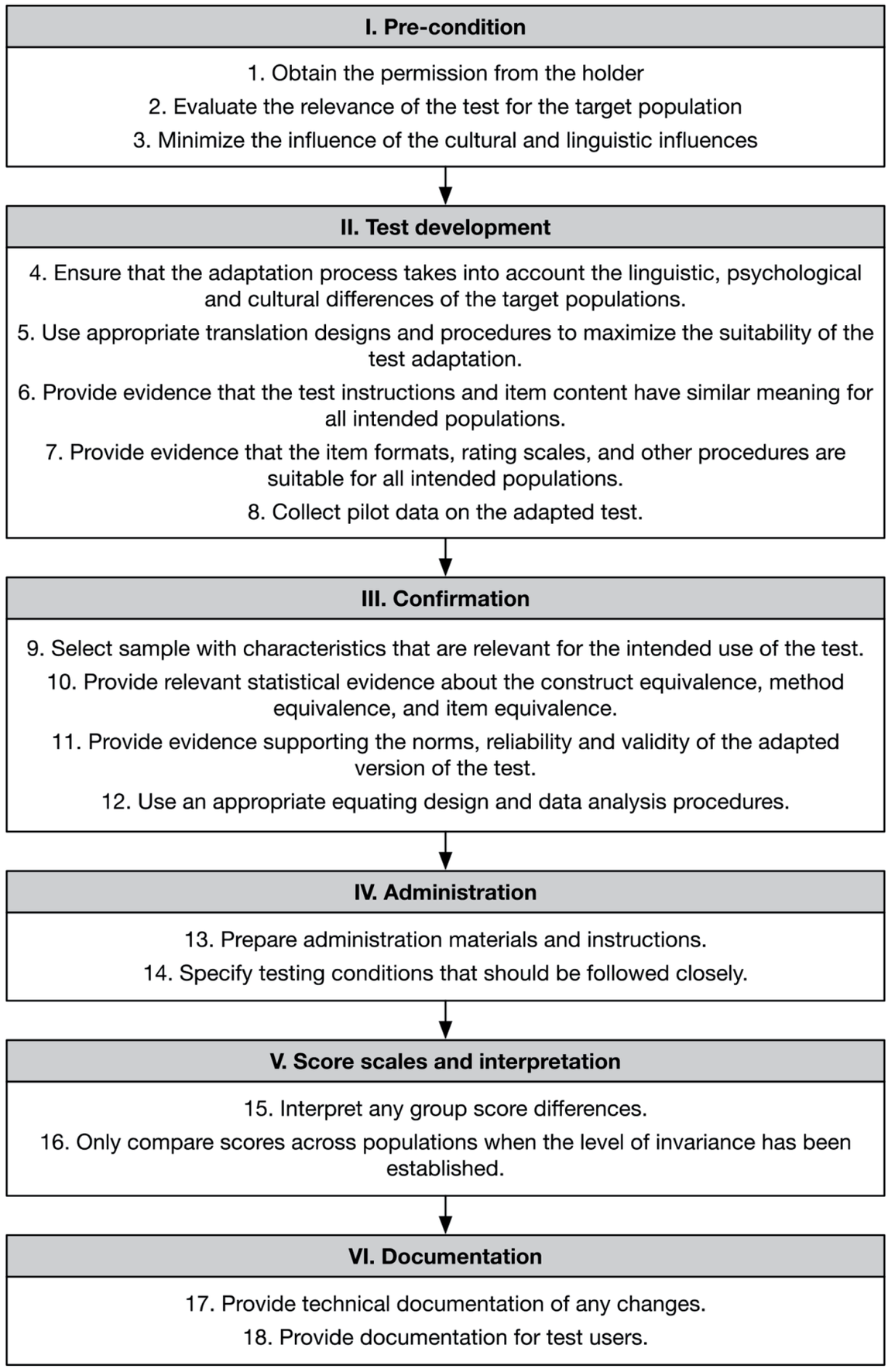

Figure 1.
Synthesis of the International Test Commission guidelines for translating and adapting tests.

interpretation'. Indeed, psychometric analyses are involved in the process of adapting items from the original language to a new language in order to ensure the quality of the translation. Failure to transfer the meaning of the original items can lead to a variation called scale error in the scale scores. As a result of this transfer of meaning, it is possible to create a structure that is different from the original scale structure. Therefore, in an adaptation study, it is necessary to ensure that the translation of the item is done correctly before starting the analysis. A consistent translation process is very important for the elimination of structural differences [6].

The most important and commonly applied psychometric analyses are presented in the following section.

3. Psychometric analyses of scale adaptations

3.1 Measuring internal consistency

According to the models of classical test theory [7], the total score (X) on a test is never fully representative of the true score (V), i.e., the exact quantity that is being measured. There is always an error (ε), so the total score is composed of the true score and the error score. Thus, we note [8]:

$$\text{Score}_{\text{Total}} = \text{Score}_{\text{True}} + \text{Score}_{\text{Error}} \quad (1)$$

The error is assumed to be random with an average of zero, so that it sometimes acts to increase the total score and sometimes to decrease it, but does not bias it in any systematic way. Since any scale has some degree of measurement error, it is never possible to determine the true score, which would be the average of all the scores a person would get if they took the test an infinite number of times [9].

The error is itself divided into two components: the random error, which is normally distributed and has a mean of 0, and the systematic error, which is asymmetrically distributed and has a mean that differs from 0. While the random error does not introduce systematic bias into the measurement, the systematic error, when it differs from 0, will cause the observed score to systematically overestimate or underestimate the true score. Thus, the true score (V) will be composed of the construct of interest (CI) and the systematic error (SE), plus the random error (RE):

$$\text{Score}_{\text{Total}} = \text{Score}_{\text{Construct of Interest}} + \text{Score}_{\text{SystematicError}} + \text{Score}_{\text{RandomError}} \quad (2)$$

Fidelity estimators are used to assess how close the observed score is to the true score.

3.1.1 Cronbach's alpha

One of the most widely used fidelity indices in the humanities and social sciences is most likely Cronbach's alpha [10]. According to Cronbach, internal consistency refers to the homogeneity of the items, i.e. how similar the test items are or, in other words, how well they measure the same dimension of a construct, i.e., its unidimensionality.

Cho and Kim [11] state that the articles by Cortina [12] and Schmitt [13] have done much to inform researchers on the use of alpha, highlighting its advantages and limitations. Other research is more radical and recommends the use of other measures of internal consistency [14, 15]. Indeed, several authors [11–13] have demonstrated that a high alpha value does not necessarily translate into homogeneity or unidimensionality of the items. Rather, alpha indicates how closely the items in a scale are related or correlated to each other.

Yet most studies of cross-cultural adaptation of scales in psychology still rely on the calculation of Cronbach's alpha as a measure of internal consistency or homogeneity; see for example [11–13]. This persistence of alpha in psychometric studies can be explained by the ubiquity of this measure since the 1950s, which allows comparisons to be made between scales. It is indeed common to rely on the alpha of the original scale to ensure the validity of a translation into a target language, by comparing the alpha of the two scales. Moreover, across research, alpha is used as a traditional benchmark for measuring internal consistency, although as we have pointed out this interpretation is biased. Sijtsma [15] finally points out that in practice it is often understood that SPSS statistical software does not offer any calculations other than homogeneity, which is of course wrong. Cho and Kim [11] conclude that alpha has become as popular as some marketing products, which are less effective than others but have a better reputation than others. They, therefore, advise authors, but also editors of scientific journals, to incorporate other indicators of internal consistency, in addition to or instead of alpha.

3.1.2 McDonald's omega

As an alternative to Cronbach's alpha, McDonald's omega [16] is the second indicator of internal consistency that is most often found in cross-cultural adaptation of scales in psychology. It is a fidelity coefficient that takes into account the strength of the association between items and a construct on the one hand, and the link between the items and the measurement error on the other. Thus, according to McDonald, the omega provides a more accurate estimate of the true reliability of the scale.

Several studies justify the use of McDonald's omega as an alternative reliability index to the alpha [14, 17, 18]. Also, some cross-cultural adaptation studies calculate the omega in addition to the alpha [19–21]. However, these studies are far from being the most representative, and none of them completely replace alpha with omega.

3.2 Factor analysis

3.2.1 Exploratory factor analysis

The main purpose of exploratory factor analysis (EFA) is to identify the underlying latent variables or factors of a measure by exploring the relationship among observed variables [22]. Roberson et al. [22] also report that, as an exploratory technique, EFA should not be used as a rigorous verification of the theoretical model; that is, in the case of cross-cultural scale adaptation, as a means of verifying the factorial adequacy of the translated scale with respect to the original scale. Finally, the authors summarise a set of good practices for conducting EFA, in terms of the statistical distribution, sample size, extraction and rotation to be applied and the matrices to be included in the publications. Comrey [23] points out in this respect that too little

information on the application of EFA is given by researchers, which makes it difficult to compare or replicate studies.

In general, EFA is used to extract latent factors from the newly translated scale. The results of this analysis are compared to the structure of the original scale to verify that the same factors are present, with a similar organisation of items within each factor. Many studies of cross-cultural scale adaptation in psychology use this process; see for example [19–21, 24].

3.2.2 Principal component analysis

Principal component analysis (PCA) is one of the most popular multivariate statistical technique in psychometric analysis in psychology. It is also likely to be the oldest multivariate technique, formalised in its current state by Hotelling [25]. According to Abdi and Williams [26], PCA analyses a data table representing observations described by several dependent variables, which are, in general, inter-correlated. Its goal is to extract the important information from the data table and to express this information as a set of new orthogonal variables called principal components. PCA also represents the pattern of similarity of the observations and the variables by displaying them as points on maps. Jolliffe [27] adds that PCA is often used to reduce the dimensionality of a data set, replacing the p variables which have been measured by a much smaller set of m components. In the case of measurement scales in psychology, p represents the items and m the factors, or dimensions, of that scale.

In cross-cultural adaptations of psychological scales, the PCA is applied instead of the EFA. The orthogonal Varimax rotation is the most common one [24], although other rotation methods are also used, but are generally not well documented [28].

Two criteria are frequently used to determine the number of factors to be extracted from the PCA. The first criterion is the widely used eigenvalue. The higher the initial eigenvalue, the more the factor explains a significant portion of the total variance. By convention, any factor with an initial eigenvalue greater than 1 is considered significant. The second criterion is Cattell's kink criterion, and it is a more stringent criterion for determining the number of factors. Here a graph displays all the points that represent the eigenvalues of the components. They are connected by a line. Only those factors that lie before the abrupt change in slope are retained. The points following this change, called the bend break, appear to form a straight horizontal line. A few publications offer eigenvalue graphs [29], but this is not common practice (**Figure 2**).

3.2.3 Confirmatory factor analysis

Confirmatory factor analysis (CFA) is a type of structural equation modelling that assesses the internal validity of an instrument or the relationships between several manifest and latent variables [30]. CFA is used to test the fit between an a priori defined theoretical model and empirically collected data. This means that the researcher must be able to specify how many factors are needed and which variables would load heavily or have near-zero loadings on each factor. Thus, on the basis of various fit indices, it is determined whether the postulated model fits the data well. When the model does not show a good enough fit, the indices exceed a threshold value, thus suggesting the rejection of the model tested.

The CFA technique is particularly well suited to cross-cultural studies. Watkins [31] states that CFA can be used to compare the equivalence of factor structures across

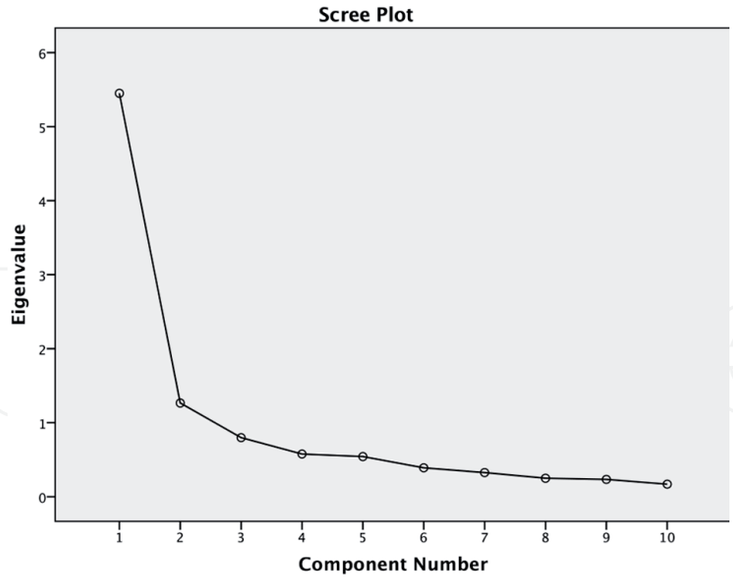

Figure 2.
Illustration of a scree plot [29].

cultures. This can be done either by collecting similar data in each culture or by collecting data in one culture and testing it against the factorial model established in the other culture. DiStefano and Hess [32] note the ubiquity of CFA in construct validation studies in psychology. Indeed, it is observed that most cross-cultural adaptations are validated or invalidated, using CFA; see for example [33–35].

The validation of the adapted scale, in comparison with the theoretical model of the original scale, necessarily relies on the consideration of fit indices, described in the next section.

3.2.4 Fit indices

The validation of the structural model calculated in the CFA is based on a set of fit indices whose thresholds indicate whether the model tested is valid or not. In other words, in the cross-cultural adaptation of psychological scales, the researcher applies the structural model of the original scale to his or her translation using a CFA in the first instance (the items are grouped into the corresponding dimensions), and then observes whether this model can be retained or should be rejected. If it is rejected, one or more other models are then applied until a satisfactory model is found, thus meeting the fit indices. This approach is applied, for example, to the French translation and validation of the Karitane Parenting Confidence Scale [20].

Several fit indices are usually calculated, of which we present here the most used among a larger set of fit indices [36], indicating the thresholds for model acceptance:

- Normed χ^2. A value below 2 suggests good model fit and below 3 acceptable model fit [37].

- Goodness of Fit Index (GFI) [38]. The value should be higher than 0.90 [39].

- Comparative fit index (CFI). The value should be higher than 0.80 [39]. CFI indicates a good model fit for values in the range between 0.95 and 1.00, whereas values in the range of 0.90 to 0.95 signify acceptable fit [37].

- Tucker-Lewis index (TLI). TLI higher than 0.95 is a commonly used cut-off criterion for the goodness of fit [40].

- Root mean square error of approximation (RMSEA). The value should be lower than 0.10 [39]. The values less than 0.05 indicate a good model fit and less than 0.08 an acceptable model fit [37].

- Standardised root means square residual (RMSR). Hu and Bentler [40] suggested the value should be lower than 0.08, while Sivo, Fan, Witta and Willse [41] suggested a cut-off value of 0.05.

- Akaike information criterion (AIC). The lower the value is, the better the fit [39].

- Normed Fit Index (NNFI). A correct model should have a value higher than 0.95 [41].

- Standardised root means square residual (SRMR). The values are expected to stay below 0.10 [37].

3.3 Convergent validity

Calculating convergent validity is an important step in measuring the validity of a scale adapted into another language. This is to ensure that the instrument really measures the construct(s) it is intended to measure and that it provides an adequate measure of the theoretical model on which it is based. A scale with good construct validity should therefore normally have high correlations with other scales measuring the same or similar constructs. Convergent correlations are therefore measured using, most often, Pearson's correlation coefficient.

In the context of cross-cultural adaptation of a scale, the translated scale is compared to one or more scales in the same language, which measure a similar psychological concept. For example, Yang, Zang, Ma et Bai [19] compared the Surgical Fear Questionnaire (SFQ) with the Hospital Anxiety and Depression Scale (HADS). The significance levels (p-value), associated with the correlation coefficients, indicate whether the links between the scales are satisfactory or not.

3.4 Time constancy

The time constancy is measured using the so-called test–retest technique. This technique consists of administering the same scale to the same subjects at two-time intervals. Generally, following the first measurement, the second measurement is carried out after 2 to 4 weeks. The scores at these two-time points are compared using a Pearson correlation coefficient, the Intraclass Correlation Coefficient (ICC) or the Kendall coefficient of concordance [42]. This technique ensures that the scale is stable over time and therefore reliable. A correlation with $0.30 < r < 0.50$ is considered as low, moderate with $0.50 < r < 0.70$ and strong with $r > 0.70$ [39].

In a cross-cultural adaptation of the Implicit Theory of Emotion Scale, Congard et al. [43] interviewed 35 subjects, 21 to 27 days apart. The Pearson correlation coefficient of 0.69 ($p < 0.001$) showed very good reliability of the scale over time.

It should be noted, however, that this technique is not relevant for certain scales in psychology, such as those measuring the perception of a product. Indeed, depending

on the use of the product, the same individual may have very different perceptions of the same product from one week to the next. This is particularly the case for scales measuring usability or user experience [29, 44, 45].

3.5 Socio-demographic analyses

The sensitivity of a cross-cultural adjustment is measured by comparisons between different modalities of the same variable. The difference in scores according to gender is often the first element of comparison. Depending on the variables and the number of modalities of the variables, researchers conventionally apply Student's t or ANOVA when there are more than two modalities.

In the adaptation of the Feelings at School (FAS) scale, Sanchez et al. [46] compared the scores between two different primary school levels (6- and 11-year-olds). The calculation of an ANOVA revealed the presence of a significant effect on the school level.

4. Conclusion

The aim of this chapter was to propose a methodological framework for psychometric analyses in the cross-cultural adaptation of psychological scales. Although the

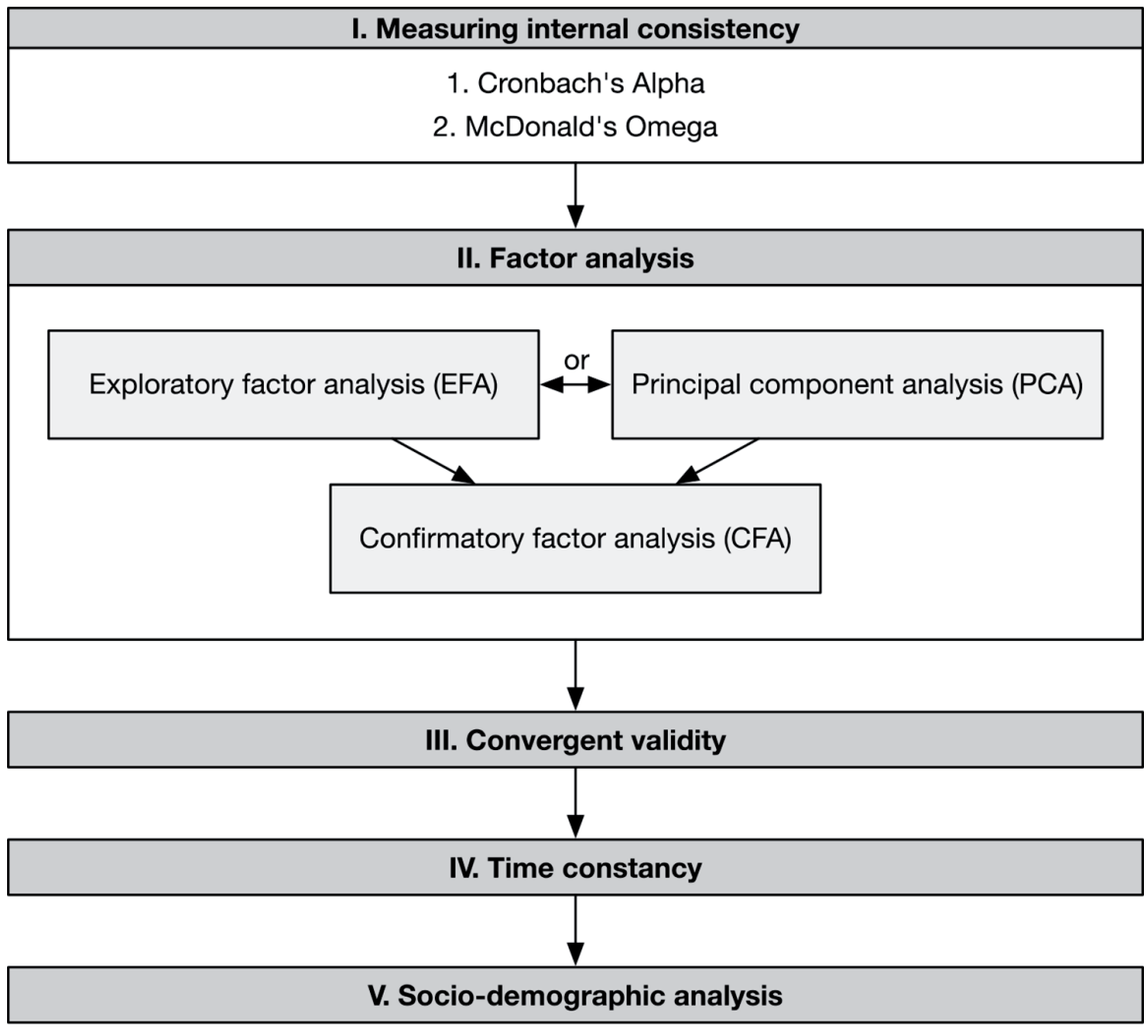

Figure 3.
Methodological framework for psychometric analysis in the transcultural adaptation of psychological scales.

choice of statistical validation tools may change from one study to another, depending on the requirements of the journal for publication or the psychometric skills of the researcher, it is possible to identify a guideline in a succession of steps that can serve as a guide to the cross-cultural adaptation of scales. This methodological line, which takes up the analyses described in this chapter, is described in **Figure 3**. It is imperfect and not exhaustive, but it will be a support that will be suitable for most of the validation of scale translations.

Author details

Guillaume Gronier
Luxembourg Institute of Science and Technology (LIST), Esch-sur-Alzette, Luxembourg

*Address all correspondence to: guillaume.gronier@list.lu

References

[1] Sousa VD, Rojjanasrirat W. Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: A clear and user-friendly guideline. Journal of Evaluation in Clinical Practice. 2011;**17**:268-274. DOI: 10.1111/j.1365-2753.2010.01434.x

[2] Gana K, Broc G, Boudouda NE, Calcagni N, Ben YS. The ITC guidelines for translating and adapting tests (second edition). Pratiques Psychologiques. 2021;**27**:175-200. DOI: 10.1016/j.prps.2020.06.005

[3] Gana K, Boudouda NE, Ben Youssef S, Calcagni N, Broc G. Transcultural adaptation of psychological tests and scales: A practical guide based on the ITC guidelines for translating and adapting tests and the standards for educational and psychological testing. Pratiques Psychologiques. 2021;**27**:223-240. DOI: 10.1016/j.prps.2021.02.001

[4] Sperber AD. Translation and validation of study instruments for cross-cultural research. Gastroenterology. 2004;**126**:124-128. DOI: 10.1053/j.gastro.2003.10.016

[5] International Test Commission. ITC Guidelines for Translating and Adapting Tests. 2nd ed. 2017. Available from: www.InTestCom.org

[6] Orçan F. Exploratory and confirmatory factor analysis: Which one to use first? Journal of Measurement and Evaluation in Education and Psychology. 2018;**9**:414-421. DOI: 10.21031/epod.394323

[7] Steyer R. Classical (psychometric) test theory. International Encyclopedia of the Social & Behavioral Sciences. 2001:1955-1962. DOI: 10.1016/B0-08-043076-7/00721-X

[8] Streiner DL. Starting at the beginning: An introduction to coefficient alpha and internal consistency. Journal of Personality Assessment. 2003;**80**:99-103. DOI: 10.1207/S15327752JPA8001_18

[9] Allen MJ, Yen WM. Introduction to Measurement Theory. Monterey: Brooks Cole; 1979

[10] Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika. 1951;**16**:297-334. DOI: 10.1007/BF02310555

[11] Cho E, Kim S. Cronbach's coefficient alpha: Well known but poorly understood. Organizational Research Methods. 2015;**18**:207-230. DOI: 10.1177/1094428114555994

[12] Cortina JM. What is coefficient alpha? An examination of theory and applications. The Journal of Applied Psychology. 1993;**78**:98-104. DOI: 10.1037/0021-9010.78.1.98

[13] Schmitt N. Uses and abuses of coefficient alpha. Psychological Assessment. 1996;**8**:350-353. DOI: 10.1037/1040-3590.8.4.350

[14] Hayes AF, Coutts JJ. Use omega rather than Cronbach's alpha for estimating reliability. Communication Methods and Measures. 2020;**14**:1-24. DOI: 10.1080/19312458.2020.1718629

[15] Sijtsma K. On the use, the misuse, and the very limited usefulness of cronbach's alpha. Psychometrika. 2009;**74**:107-120. DOI: 10.1007/s11336-008-9101-0

[16] McDonald RP. Generalizability in factorable domains: "Domain validity and generalizability". Educational and Psychological Measurement. 1978;**38**:75-79. DOI: 10.1177/001316447803800111

[17] Revelle W, Zinbarg RE. Coefficients alpha, beta, omega, and the glb: Comments on sijtsma. Psychometrika. 2009;**74**:145-154. DOI: 10.1007/s11336-008-9102-z

[18] Kelley K. Supplemental material for confidence intervals for population reliability coefficients: Evaluation of methods, recommendations, and software for composite measures. Psychological Methods. 2016;**21**:69-92. DOI: 10.1037/a0040086.supp

[19] Yang G, Zang X, Ma X, Bai P. Translation, cross-cultural adaptation, and psychometric properties of the Chinese version of the surgical fear questionnaire. Journal of PeriAnesthesia Nursing. 2022;**000**:1-7. DOI: 10.1016/j.jopan.2021.08.004

[20] Ribeiro MH, Coulon N, Guerrien A. French translation and validation of the Karitane parenting confidence scale. European Review of Applied Psychology. 2022;**72**:1-8. DOI: 10.1016/j.erap.2022.100759

[21] Villarreal-Zegarra D, Torres-Puente R, Otazú-Alfaro S, Al-kassab-Córdova A, Rey de Castro J, Mezones-Holguín E. Spanish version of Jenkins sleep scale in physicians and nurses: Psychometric properties from a Peruvian nationally representative sample. Journal of Psychosomatic Research. 2022;**157**:110759. DOI: 10.1016/j.jpsychores.2022.110759

[22] Roberson RB, Elliott TR, Chang JE, Hill JN. Exploratory factor analysis in rehabilitation psychology: A content analysis. Rehabilitation Psychology. 2014;**59**:429-438. DOI: 10.1037/a0037899

[23] Comrey AL. Common methodological problems in factor analytic studies. Journal of Consulting and Clinical Psychology. 1978;**46**:648-659. DOI: 10.1037/0022-006X.46.4.648

[24] Bled C, Bouvet L. Validation of the French version of the object spatial imagery and verbal questionnaire. European Review of Applied Psychology. 2021;**71**:100687. DOI: 10.1016/j.erap.2021.100687

[25] Hotelling H. Analysis of a complex of statistical variables into principal components. Journal of Education & Psychology. 1933;**24**:417-441. DOI: 10.1037/h0071325

[26] Abdi H, Williams LJ. Principal component analysis. Wiley Interdisciplinary Reviews: Computational Statistics. 2010;**2**:433-459. DOI: 10.1002/wics.101

[27] Jolliffe IT. Rotation of principal components: Choice of normalization constraints. Journal of Applied Statistics. 1995;**22**:29-35. DOI: 10.1080/757584395

[28] Henry A, Gagnon J. Psychometrics properties of the francophone version of the displaced aggression questionnaire. Pratiques Psychologiques. 2022;**28**:29-42. DOI: 10.1016/j.prps.2021.04.002

[29] Gronier G, Baudet A. Psychometric evaluation of the F-SUS: Creation and validation of the French version of the system usability scale. International Journal of Human Computer Interaction. 2021;**37**:1571-1582. DOI: 10.1080/10447318.2021.1898828

[30] Byrne BM. Structural equation Modeling with AMOS, EQS, and LISREL: Comparative approaches to testing for the factorial validity of a measuring instrument. International Journal of

Testing. 2001;**1**:55-86. DOI: 10.1207/s15327574ijt0101_4

[31] Watkins D. The role of confirmatory factor analysis in cross-cultural research. International Journal of Psychology. 1989;**24**:685-701. DOI: 10.1080/00207598908247839

[32] DiStefano C, Hess B. Using confirmatory factor analysis for construct validation: An empirical review. Journal of Psychoeducational Assessment. 2005;**23**:225-241. DOI: 10.1177/073428290502300303

[33] Fung SF. Validity of the brief resilience scale and brief resilient coping scale in a Chinese sample. International Journal of Environmental Research and Public Health. 2020;**17**:1-9. DOI: 10.3390/ijerph17041265

[34] Sahlan RN, Todd J, Swami V. Psychometric properties of a Farsi translation of the functionality appreciation scale (FAS) in Iranian adolescents. Body Image. 2022;**41**:163-171. DOI: 10.1016/j.bodyim.2022.02.011

[35] Petot D, Petot JM, Fouques D. Factor structure and psychometric properties of the Achenbach and Rescorla's youth self-report French adaptation. European Review of Applied Psychology-Revue Europeenne de Psychologie Appliquee. 2022;**72**:1-13. DOI: 10.1016/j.erap.2021.100701

[36] Hu L, Bentler PM. Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. Psychological Methods. 1998;**3**:424-453. DOI: 10.1037//1082-989x.3.4.424

[37] Schweizer K. Some guidelines concerning the modeling of traits and abilities in test construction. European Journal of Psychological Assessment. 2010;**26**:1-2. DOI: 10.1027/1015-5759/a000001

[38] Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. Structural Equation Modeling. 2002;**9**:233-255. DOI: 10.1207/S15328007SEM0902_5

[39] Tabachnick BG, Fidell LS. Using Multivariate Statistics. seventh ed. Boston: Pearson; 2019

[40] Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling. 1999;**6**:1-55. DOI: 10.1080/10705519909540118

[41] Sivo SA, Xitao FAN, Witta EL, Willse JT. The search for "optimal" cutoff properties: Fit index criteria in structural equation modeling. The Journal of Experimental Education. 2006;**74**:267-288. DOI: 10.3200/JEXE.74.3.267-288

[42] Noble S, Scheinost D, Constable RT. A decade of test-retest reliability of functional connectivity: A systematic review and meta-analysis. NeuroImage. 2019;**203**:116157. DOI: 10.1016/j.neuroimage.2019.116157

[43] Congard A, Le Vigouroux S, Antoine P, Andreotti E, Perret P. Psychometric properties of a French version of the implicit theories of emotion scale. European Review of Applied Psychology. 2022;**72**:100728. DOI: 10.1016/j.erap.2021.100728

[44] Erdinç O, Lewis JR. Psychometric evaluation of the T-CSUQ: The Turkish version of the computer system usability questionnaire. International Journal of Human Computer Interaction. 2013;**29**:319-326. DOI: 10.1080/10447318.2012.711702

[45] Lallemand C, Koenig V, Gronier G, Martin R. Création et validation d'une version française du questionnaire AttrakDiff pour l'évaluation de l'expérience utilisateur des systèmes interactifs. European Review of Applied Psychology. 2015;**65**:239-252. DOI: 10.1016/j.erap.2015.08.002

[46] Sanchez C, Baussard L, Blanc N. Validation of the French and enriched version of the feelings about school (FAS) with students aged 6 to 11. Psychologie Française. 2022:1-17. DOI: 10.1016/j.psfr.2021.12.001