Procedures for Translation and Validation of the Instrument

Binoy Thomas, PhD Scholar, VIT Business School, Vellore Institute of Technology,Vellore, India.
P. Subhashree, Corresponding Author, Professor & Dean, VIT Business School, Vellore Institute of Technology, Vellore, India.

Abstract:

The precision in the translation of the survey instrument is a key methodological issue, because the quality of translation process has a crucial role in determining the accuracy and quality of the research study. The errors in translation may distort the original purpose of the instrument itself. In this context, researchers have recommended for a comprehensive process of translation. The present study is a literature review paper on the standardised procedures to be followed for translating a survey instrument from the original language into the vernacular language and for validating the translated questionnaire. The study has conducted a systematic review on 139 articles which discusses the instrument translation and validation procedures. The present study would contribute a review on existing literature on instrument translation and validation.

Keywords: *Forward-back translation, Instrument translation, Instrument validation*

I. INTRODUCTION

Translation is of critical importance in the development of a cross-cultural survey instrument for data collection. Eventhough the starting point for any survey research is the conceptual framework under study, the vehicle which drives the research is the survey instrument (questionnaire). The translated instrument must have equivalent meaning in each context, devoid of cultural bias, and can be clearly comprehended by all respondents. In the researches at international level, the meaning of the constructs shall be equivalent across diverse contexts. If not, the results cannot be compared across different linguistic contexts; inappropriate conclusions may be drawn and the value of the research will be negated.

What is the necessity for translating a survey instrument? Most of the established scales and survey instruments are originally prepared in English language. The questionnaire prepared in English language is difficult to follow for those are not well versed in English language. If the respondents are not comfortable with the questionnaire prepared in English language, there is a higher probability for rejecting the questionnaire or for giving faulty responses. To handle such issues, translating the questionnaire into the mother tongue (local) language of the respondents is a necessity for ensuring better acceptability among the respondents.

Moreover, the data collection at international level would be increasingly difficult for the respondents who live different cultural contexts and language backgrounds. Such situations mandate to translate the original instruments into different languages and cultural contexts. Also, the process of translating and validating an instrument is faster and cheaper, while compared to developing a new survey instrument.

The precision in the translation of the survey instrument is a key methodological issue [1],

because the quality of translation process has a crucial role in determining the accuracy and quality of the research study [2]. The errors in translation may distort the original purpose of the instrument itself. In this context, Wang [3] stated that the simple translation and use of a questionnaire in another language context is inappropriate. So he recommend for a comprehensive process of translation.

The translation and adaption of the research instrument is to be necessarily followed by an appropriate validity test; else it would effect in erroneous interpretations of the results. In order to use a translated instrument in cross-cultural studies, it is to be validated in the new context.

The research question of the study is: What are the various procedures to be followed for translating and validating a survey instrument? The research objective of the study is to sketch a comprehensive review of literature on the instrument translation and validation.

II. INSTRUMENT TRANSLATION

Translation is the process of transferring ideas and thoughts from the one language (source language) into another language (foreign or target language) either in written form or in oral form [4]. The translation process includes not only the change of the language, but also the transfer of the subjective experiences expressed in the original language into the targeted language [5].

There is no any universal agreement on how to adapt a survey instrument for using in another cultural and language context [6]. But researchers have proposed various methods and guidelines for the same. Each item in the original questionnaire is to be properly adjusted into the cultural context of the targeted respondents [7]. Moreover, the

translation process is to be mandatorily followed by instrument validation, since the translation might alter the characteristics of the original instrument [8].

Researchers [9], [10] have generally proposed three methods for translating survey instruments from original language into a foreign language: (i) One-shot translation (forward translation), (ii) Forward-back translation, and (iii) Simultaneous development or translation. One shot translation is unidirectional; it translates the survey instrument from the original language to the target language. The Forward-back translation is a bi-directional process; it starts with translating the instrument from the original language into the target language, followed by back-translating the translated instrument into the original language, and finally by comparing both the original instrument and the back-translated instrument [4].

A systematic review of literature was conducted to identify various procedures followed for instrument translation in the previous literature. 139 peer reviewed research articles were identified and collected from various databases and search engines: EBSCOhost (22), Emerald Insight (22), Google Scholar (29), SAGE Journals (9), Science Direct (12), Scopus (23), Wiley Online Library (22). Keywords such as 'Translation' AND 'Questionnaire' OR 'Survey Instrument' were used for identifying the articles. Time span of 20 years (from 2000 to 2019) was chosen for identifying the research articles. The relevant articles were shortlisted reading the abstract of the articles. The details regarding the database and the year of publication are as shown in Table I.

Table I: Database & Year of Publication of Articles

Years	EBSCO--host	Emerald Insight	Google Scholar	SAGE	Science Direct	Scopus	Wiley	Total
2000	1							01
2001	1		2	1			3	07
2002	1		1			1	1	04
2003	3	1		1	3		1	09
2004			4		1	1	2	08
2005			3	1			2	06
2006	3	2	4	1		2	1	13
2007	4	3	3	2	1	3	3	19
2008					2		1	03
2009	1	1	1			3	1	07
2010	1	1	2	2	1	4	3	14
2011	1	2			1		1	05
2012	1	3	1					05
2013		1			2	2		05
2014	1	1	2		1	3	1	09
2015	2		3	1		1	1	08
2016		1	1				1	03
2017	1	2	1			3		07
2018	1	2						03
2019		2	1					03
Total	22	22	29	09	12	23	22	139

All these 139 articles were reviewed by authors and identified various translation methods used in them (Fig. I).
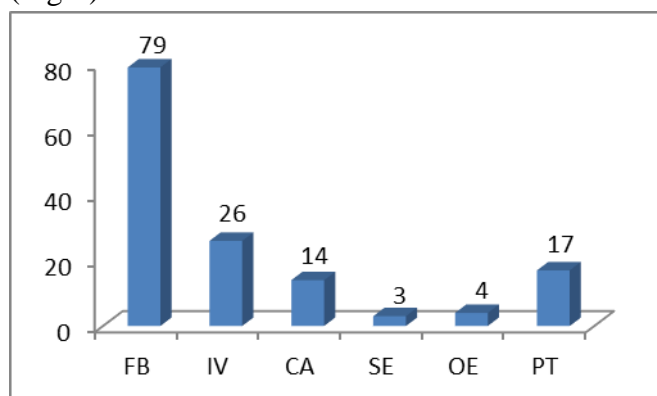

Fig. I. Translation Methods used in the Previous Literature

Among these 139 articles, Forward-Back translation (FB) was used in 57% (79) of the articles. In 26 articles, Independent versions (IV) were undertaken, and Committee Approach (CA) in 14 articles. Subjective Evaluation (SE) and Objective Evaluation (OE) were used in three and four articles respectively. 17 articles used Pre-testing (PT) of the questionnaire before carrying out the final study.

Forward-back translation proposed by [4] is based on two assumptions: (i) The cultural differences of the respondents will influence the response given on the questionnaire; (ii) The language of the questionnaire influences the research process, even if the language is the only difference between the two

cultures, i.e. originally developed, and currently used. The back-translation process ensures the accuracy of the translation, equivalence of the meanings in the original and translated instruments, and increases the face validity of the measurement items. The forward-back translation model consists of 4 steps.

A. Forward Translation

The survey instrument is translated from the original language into the language of the targeted respondents (vernacular, local or foreign language). This process is independently performed by at least two bilingual translators. Chen et al [2] instructed the translators to retain the form and meaning of the items as close to the original questionnaire, and to give priority for the meaning over the sentence form.

B. Synthesized Forward Translation

Independent and parallel versions of the survey instrument are synthesized into a single version by a third translator. In the case of discrepancies or disagreements, the items are revised until a consensus is reached [2]. The two translations were then compared to assess the item-by-item consistency.

C. Back Translation

The synthesized translated instrument is back-translated into the language of the original instrument by two independent bilingual translators. These translators are not given the questionnaire in the original language. Back translation ensures the consistency and equivalence of the meaning between the original instrument and the translated instrument [11].

D. Synthesized Back Translation

The two independent versions of the back translated instruments are reviewed by a third back-translator. Evaluating the differences among the back-translated versions, a synthesized back-translated version of the instrument is derived.

Researchers [11], [12] have stated that the primary emphasis in the translation process is on achieving 'meaning equivalence' and 'functional equivalence', rather than a mere identical word-by word literal translation. The forward-back translation process is repeated until the two versions achieve the equivalence. Multiple iterations though appear to be complex and time consuming ensures that the translated instrument is appropriate for the new context.

Based on the above mentioned forward-back translation procedures, various authors have translated the original English instruments into other foreign languages: Thai [2], Portuguese [11], Chinese [12], German [13], Ukrainian [14], Norwegian [15], etc.

Some researchers [2], [3] specify that the translators need to be fluent in the target language and have a good understanding of the original language. The translators were instructed to state the difficulties they faced while translating the words, idioms or phrases [31].

Expert Committee Review: An expert committee which consist of translators, linguists, psychometricians, and practitioners review the synthesized translated version and the synthesized back-translated version. They evaluates whether both versions of instrument reflect same meaning, and the translated instrument is functional in the targeted culture [3]. Two forms of collaborative-based work, i.e. the committee approach and the expert team approach, are proposed by[16].

To ensure the meaning equivalence of the items in the original version and the back translated version of the questionnaire, researchers [2], [11], [12] suggest to extend the forward-back translation process by Subjective evaluation; (ii) Objective evaluation; and (iii) Pre-testing and pilot study.

(i) Subjective evaluation: In the subjective evaluation, a qualitative approach is used for assessing the meaning equivalence of the two versions (original and back translated versions) of the questionnaire. If substantial differences in meaning were found between items, such items are

again undergone the translation-back-translation-review process until no items exhibiting substantial differences in meaning emerged.

(ii) Objective evaluation: In the objective evaluation, a quantitative evaluation on the meaning equivalence of the original version and the back-translated version of the questionnaire are conducted. The Objective evaluation is made based on two criteria: (i) Comparability of language, and (ii) Similarity of interpretability [17]. Comparability of language discusses the similarity of words and phrases sentences used in two versions; Similarity of interpretability evaluates the similarity in the meaning of items. These two criteria is assessed on a 7 Point Likert Scale [17], where Score 1 represents 'Extremely comparable' or 'Extremely similar', and Score 7 represents 'Not at all Comparable' or 'Not at all similar'. Any items with mean score greater than 3.0 are flagged as potentially problematic and are undergone further investigations.

(iii) Pre-testing and Pilot test: Pre-testing and pilot test of the translated questionnaire for reducing translator bias and improving functionality of the questionnaire [4]. Pilot test ensures the quality and dependability of the translated instrument [2]. Researchers [10] - [12] also proposed to conduct pilot study before final study in order to collect comments on the interpretability of the instrument.

Various methods of equivalence are proposed by [31] for establishing cross-cultural equivalence in the process of instrument translation; they are: content equivalence, semantic equivalence, technical equivalence, criterion equivalence, conceptual equivalence.

III. INSTRUMENT VALIDATION

If the survey instrument (questionnaire) is originally developed in a context different to that of the present study, it is to be validated before applying for the present study. The validity of a questionnaire, in the strict sense, is specific to the particular population in which it is originally developed. The instrument validation is mandatory if the questionnaire is modified at least partially or is translated to another language [1], [5], [8]. If the empirical research is conducted without validating the questionnaire, the results cannot be generalised than to the sample tested in the research study [18]. Also the instrument validation reduces the measurement errors [19].

Based on this psychometric method, researchers [18], [19], [20] suggested various procedures for instrument development and validation. Generally there are 6 steps for the instrument validation:

A. Establish Content Validity

Content validity refers to the degree to which the items in the questionnaire theoretically capture and sufficiently cover all aspects of the variables under study [21], [22]. The content validity assesses whether the questionnaire neglects any construct-relevant items, includes any construct-irrelevant items. Moreover it also testifies whether the questionnaire is well suitable for psychometric testing [23]. The content validity is assessed by a panel of experts; it can be minimum three [24] or in a range of two to twenty [25].

The experts evaluate each item based on four criteria: (i) Representativeness, (ii) Clarity, (iii) Factorial structure, and (iv) Comprehensiveness [26]. Representativeness refers to the ability of the individual items to represent the content domain; Clarity evaluates how clearly the items are worded; the Factor structure measures whether the items are properly related to the corresponding factors (constructs). Based on the scored assigned to Representativeness and Clarity, the Content Validity Index (CVI) for each item is calculated. Based on the assignment of the Factorial Structure by the experts, the Factorial Validity Index (FVI) is assessed. The items with lower scores are either deleted or reviewed based on the evaluative comments of the experts on the 'comprehensiveness'. Based on the content validation, the draft questionnaire is revised as long as items are seen as the most appropriate.

Two-stage Delphos' technique for content validation is proposed by [27]. In the first step, the experts are asked to comment whether each item is appropriate to be part of the questionnaire; in the second step, they are again contacted to verify whether the suggested modifications have been incorporated properly.

B. Run a Pre-testing and Pilot Test

Researchers [19], [28], [29] have proposed to conduct the pre-test of the questionnaire among the experts in order to assess the clarity in wording, sequence of questions, relevance, and usefulness of each item. Pretesting helps the researcher to identify the errors such as double barrelled items, the tautological problems, ambiguities, etc. Once the errors are corrected, the pilot study is conducted among a sub-set of the respondents in order to collect quantitative data for further statistical analysis. Pilot test is recommended by [29] to assess the initial validity of the survey instrument, to purify the instrument, and to enrich the clarity of the items and instructions.

C. Clean the Collected Data

After collecting pilot data, the data is to be cleaned. The negatively phrased questions are to be coded reversely. The data entry errors are to be rectified and the missing values are to be replaced. Multi-collinearity among the measurement items is to be checked. If the inter-item correlations are greater than 0.9, there is high possibility for multi-collinearity. If there are couples of items too correlated, only one item among them is to be retained.

D. Ensure Construct Validity

Construct validity is confirmed through convergent validity, and discriminant validity. The items of a construct shall strongly correlate to the construct to which they are theoretically associated, and weakly correlated to all other constructs. Eventhough there are criteria validity, concurrent validity and predictive validity, etc. in the instrument validation procedures, they are very rarely used [30].

Factor analysis can be used to establish construct validity. Researchers [2], [12] examined the factor structure of the translated questionnaires. In the instrument development stage, Exploratory Factor Analysis is conducted, whereas Confirmatory factor analysis is conducted in the context of instrument validation. Goodness of Fit can be assessed through various model fit indices. The construct validity can also be assessed through Rasch model, a model based on Item Responses Theory. Convergent validity can be measured also through Item Reliability, Composite reliability, Average Variance Extracted, etc. The discriminant validity is assessed by comparing the Cronbach's Alpha value of the construct to its correlations with other constructs in the model.

E. Check Internal Consistency (Reliability)

Internal consistency deals with the homogeneity of the responses given to the items within a scale. Reliability of a construct is assessed through Cronbach's Alpha Coefficient, Composite Reliability, Split-Half, Item - Construct (Total Item) Correlation, Test - Retest, etc. Cronbach's alpha coefficient is the most commonly used internal consistency measure. Measuring the Composite Reliability overcomes the flaws in the Cronbach's Alpha. Test - Retest reliability proves the consistency of the instrument over time. An Item-Total Item Correlation matrix measures whether the correlation coefficient of the items that are assigned to the constructs is higher that of the items not assigned to the constructs.

F. Revise the Survey Instrument

The survey instrument is revised or deleted based on the information gathered from the previous stages. If major changes are made (removal of substantial amount of questions, inclusion of many new questions, etc.), further pilot test and the subsequent procedures are to be performed. If items have cross loadings or poor goodness of fit indices, the items are either deleted or rephrased (if it is an important item for the study) to improve data-model fit.

However, any questionnaire can never be fully validated; validity of any instrument is limited to a specific population, under certain defined conditions. Each type of validity is distinct; a questionnaire can have one type of validity, but not another.

IV. CONCLUSION

After a detailed review of literature, the authors have finally reached on what all procedures (Figure-II) are to be followed for carrying out translation and validation in the context of cross-sectional studies. Thus, the present study has contributed detailed a review of literature on instrument translation and validation. It would give a solid foundation for further empirical researches.
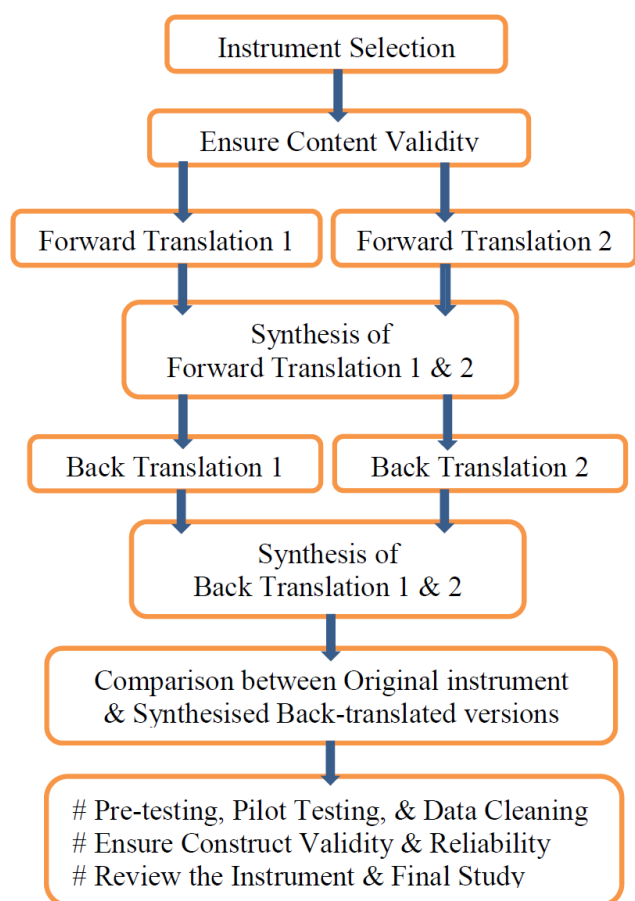

Fig. II. Procedure for Instrument Translation & Validation

The researchers have pointed out various strategies to remove diverse forms of bias and limitations inherent in the use of an instrument designed in one culture for use in another. There is still scope for future research in the area of instrument translation and validation; a deeper exploration of the literature review will help to mould more rigorous approaches for the instrument translation and validation.

REFERENES:

1. O.J. Kvamme, J. Mainz, A. Helin, M. Ribacke, F. Olesen, and P. Hjortdahl, "Interpretation of questionnaires: a translation method problem," in Norwegian Medicine, vol. 113 (10), 1998, pp. 363-366.
2. H.C. Chen, Cross-cultural Construct Validation of the Learning Transfer System Inventory in Tawian, Doctoral Dissertations 2254, Louisiana State University, 2003.
3. W. Wang, H. Lee, and S.J. Fetzer, "Challenges and strategies of instrument translation," in Western Journal of Nursing Research, 2006, p. 310.
4. R.W. Brislin, "The wording and translation of research instruments", in W.J. Lonner, and J.W. Berry (Eds.), Cross-cultural Research and Methodology Series, Vol.8, Field Methods in Cross-cultural Research, CA: Sage Publications Inc., 1986, pp. 137-164.
5. R. Fitzpatrick, C. Davey, M.J. Buxton, and D.R. Jones, "Evaluating patient-based outcome measures for use in clinical trials," in Health Technology Assessment, vol. 2 (14), 1998.
6. O. Bordeianu, and L. Morosan-Danila, "Development and validation of research instruments for cross-cultural studies in economics and management," in Proceedings of 20th International Economic Conference, May 2013.
7. L. Gjersing, R.M. Caplehorn, and T. Clausen, "Cross-cultural adaptation of research instruments: language, setting, time and statistical considerations," BMC Medical Research Methodology, vol. 10 (13), 2010.
8. A. Fraser, B.C. Delaney, A.C. Ford, M. Qume, and P. Moayyedi, "The short-form Leeds Dyspepsia questionnaire validation study," in Alimentary Pharmacology & Therapeutics, vol. 25, 2009, pp. 477-486.
9. C.H. Hui, and H.C. Triandis, "Measurement in cross-cultural psychology: a review and comparison of strategies," in Journal of Cross-Cultural Psychology, vol. 16 (2), 1985, pp. 131-152.
10. M. Bullinger, R. Anderson, D. Cella, and N. Aaronson, "Developing and evaluating cross-cultural

instruments from minimum requirements to optimal models," in Quality of Life Research, vol.2 (6), 1993, pp. 451-459.

11. R. Velada, A. Caetano, R. Bates, and E. Holton, "Learning transfer - validation of the learning transfer system inventory in Portugal," in Journal of European Industrial Training, vol. 33 (7), 2009, pp. 635-656.

12. S. Khasawneh, R.A. Bates, and E.F. Holton, "Construct Validation of an Arabic version of the Learning Transfer System Inventory (LTSI) for use in Jordan," in International Journal of Training and Development, vol. 10 (3), 2006, pp. 180-194.

13. R. Bates, S. Kauffeld, and E.F. Holten, "Factor structure and predictive validity of the German version of the learning transfer systems inventory," in Journal of European Industrial Training, vol.31 (3), 2007, pp. 195-211.

14. B.V. Yamkovenko, E.H. Ill, and R.A. Bates, "The Learning Transfer System Inventory (LTSI) in Ukraine: the cross-cultural validation of the instrument," in Journal of European Industrial Training, vol. 31 (5), 2007, pp. 377-401.

15. I.B. Erdvik, N.C. Overby, and T. Haugen, T. "Translating, reliability testing, and validating a Norwegian questionnaire to assess adolescents intentions to be physically active after high school graduation," in SAGE Open, 2015, pp. 1-6.

16. S.P. Douglas, and C.S. Craig, "Collaborative and iterative translation: an alternative approach to back translation," in Journal of International Marketing, vol. 15 (1), 2007, pp. 30-43.

17. A.D. Sperber, R.F. DeVellis, and B. Boehlecke, "Cross-cultural translation: methodology and validation," in Journal of Cross-Cultural Psychology, vol. 25, 1994, pp. 501-524.

18. B.B. Flynn, S. Sakakibara, R.G. Schroeder, K.A. Bates, and E.J. Flynn, "Empirical research methods in operations management," in Journal of Operations Management, vol. 9 (2), 1990, pp. 250-284.

19. M.K. Malhotra & V. Grover, "An assessment of survey research in POM: from constructs to theory," in Journal of Operations Management, vol. 16, 1998, pp. 407-425.

20. J.V. Saraph, P.G. Benson, and R.G. Schroeder, "An instrument for measuring the critical factors of quality management," in Decision Sciences, vol. 20 (4), 1989, pp. 810-829.

21. J.D. Brown, Testing in Language Programs, Upper Saddle River, NJ: Prentice Hall Regents, 1996.

22. S.N. Haynes, D.C. Richard, and E.S. Kubany, "Content validity in psychological assessment: a functional approach to concepts and methods," in Psychological Assessment, vol.7 (3), 1995, pp. 238-247.

23. L. Davis, "Instrument review: getting the most from your panel of experts," in Applied Nursing Research, vol. 5, 1992, pp. 194-197.

24. M. Lynn, "Determination and quantification of content validity," in Nursing Research, vol. 35, 1986, pp. 382-385.

25. R.K. Gable, and M.B. Wolf, M.B. Instrument Development in the Affective Domain Measuring Attitudes and Values in Corporate and School Settings, New York: Kluwer Academic Publishers, 1993.

26. D.M. Rubio, M. Berg-Weger, S.S. Tebb, E.S. Lee, and S. Rauch, "Objectifying content validity: conducting a content validity study in social work research," in Social Work Research, vol. 27 (2), 2003, pp. 94-104.

27. H.C. Aguilar, P.P. Cortes, A.V. Guerrero, C.D. Herrera, and J.A. Orozco, "Construction and validation of an instrument to evaluate critical reading of research papers," IN Proceedings from ICERI Conference 2011, Madrid, Spain.

28. P.J. Singh, and A. Smith, "An empirically validated quality management measurement instrument," in Benchmarking: An International Journal, vol. 13 (4), 2006, pp. 493-522.

29. R. Ahmed, and N.A. Mohamed, "Development and validation of an instrument for multidimensional top management support," in International Journal of Productivity and Performance Management, vol. 66 (7), 2017, pp. 873-895.

30. T.E. Elliott, R.R. Regal, B.A. Elliott, and C.M. Renier, "Design and validation of instruments to measure knowledge," in Journal of Cancer Education, vol. 16 (3), 2001, pp. 157-162.

31. J. Smit, C.E. van den Berg, L.G. Bekker, S. Seedat, D.J. Stein, "Translation and cross-cultural adaptation of a mental health battery in an African setting," in African Health Sciences, vol. 6 (4), 2006, pp. 215-222.