

Luella Liçi, 2022

Volume 7 Issue 3, pp.154-168

Received: 04th September 2021

Revised: 01st December 2021, 10th January 2022

Accepted: 11th January 2022

Date of Publication: 19th January 2022

DOI- <https://doi.org/10.20319/pijss.2022.154168>

This paper can be cited as Lici, L. (2022) Translation Quality in International Large-Scale Assessments in Education. *PEOPLE: International Journal of Social Sciences*, 7(3), 154-168.

This work is licensed under the Creative Commons Attribution-Non-commercial 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc/4.0/> or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA

TRANSLATION QUALITY IN INTERNATIONAL LARGE-SCALE ASSESSMENTS IN EDUCATION

Luella Liçi

PhD Candidate, Department of English Language, Faculty of Foreign Languages, University of Tirana, Tirana, Albania
luella_lic@yahoo.com

Abstract

During the last 20 years, large-scale international assessments in education became part of international research. The increased number of countries participating in the international studies of PISA, TIMSS, and PIRLS and translation of their assessment instruments in many languages of the world is accompanied by the request to ensure translation quality. Translation quality and its management became and are an important concern for the PISA, TIMMS, and PIRLS. The linguistics quality control centres have developed translation policies, standards, and procedures. This paper aims to present the translation policy and translation quality procedures implemented by PISA, TIMSS, and PIRLS in the last 20-years. The paper is focused on the need for translation quality and illustrates how translation quality is implemented in practice. The paper is based on the translation guidelines, secondary data from the technical reports of these international assessment studies, and recent research findings. This paper concludes that the suitable policy and procedures of translation quality used by PISA, TIMSS, and PIRLS have

provided a high quality of translated instruments that has enabled the comparison of assessment results. The article also highlights PISA's, TIMSS's, and PIRLS's contribution to the development of translation quality theory and practice of the educational materials.

Keywords

Translation Quality, Translation Quality Assessment, Translation Policy, Translation Quality Standards, Adaptation, Verification

1. Introduction

The increased number of countries participating in international studies such as PISA, TIMSS, and PIRLS and the translation of their assessment instruments in many languages of the world is accompanied by the request to ensure translation quality.

Translation quality and its management is an important concern of PISA, TIMSS, and PIRLS. Maxwell (1996) reported that “Because of the inherent risk of error or inequity in the translations was obvious, translation validity was an issue from the very beginnings of TIMSS” (p. 8-1). Grisay et al (2007) noted that providing linguistic and cultural equivalence through the different national versions of the assessment instruments has become an increasingly critical challenge. Kelly and Malak (2003) wrote that so that to make well-grounded comparisons, it is significant to ensure equivalence of the passages and items across languages. Dept et al. (2017) noted that equivalence is not only a scope – it is an underlying requirement in large-scale, cross-linguistic, cross-national, and cross-cultural comparative assessments, which fulfilment guarantees cross-cultural comparison is invalidated.

Aiming to ensure the translation quality, PISA, TIMSS, and PIRLS have developed and implemented special translation quality policies and procedures. This paper intends to present an overall picture of translation quality implemented by PISA, TIMSS, and PIRLS, including translation policy, translation standards, and translation/adaptation procedures. This paper may be of interest to two categories of specialists: those who are interested in translation quality issues and those who are interested to understand how the translation quality is implemented by PISA, TIMSS, and PIRLS.

2. Literature Review

Below are presented the contributions that have most influenced translation quality policy and practice implemented by PISA, TIMSS, and PIRLS.

Translation quality assessment (TQA) has become a crucial issue in translation studies. Many researchers accept that it is a difficult, complicated, delicate, and problematic issue.

Yet a special contribution has been made in translation quality assessment. The product of this contribution consists of translation quality assessment theoretical approaches and models. Some of these contributions have greatly influenced the translation quality policy and the practice implemented by PISA, TIMSS, and PIRLS.

Reiss (2000) has developed a qualitative model, Williams (2009) and some others have presented quantitative models, Nord (1997), House (1977, 1981, 1997, 2015) and some others submitted integrated models. Although translation quality assessment models differ from each other, most of them focus on the translated product.

House's TQA model (1977, 1981, 1997, 2015) is a leading model in the field of translation quality assessment. House (2015) explained that it is an eclectic model. For the translation quality assessment, this model provides comparative analyses of the source and target text based on three levels: language/text, register (field, tenor, and mode), and genre. House introduces a translation typology composed of overt translation and covert translation.

Williams (2004) argues classification of all types and approaches to translation quality assessment into two categories: a) quantitative models that deal with micro textual analysis of the texts and errors in translation, and b) qualitative models that deal with a macro textual analysis of the texts and do not concentrate on finding the number of errors.

The translation and adaptation, a field of study and practice, is developed at a rapid pace in the last 20 years thanks to the contributions of some number of researchers such as Byrne, (2008); Cook and Schmitt-Cascallar (2005); Grégoire and Hambleton (2009); Hambleton, Merenda, and Spielberger (2005); Hambleton and Patsula (1998); Harkness (2007); Harkness, Edwards, et al. (2010); Harkness, Villar, and Edwards (2010); Harkness van de Vijver, and Johnson (2003); Iliescu (2017); Nida (1991); Rios and Sireci (2014); Sireci (2005); van de Vijver and Leung, (1997), Vinay and Darbelnet (1995). Technical advances have been made in the area of the methodology of test and questionnaire translation/adaptation and translation quality standards and quality metrics methodology.

Nida and Taber (1969, 1982) link adaptation with cultural translation. According to Nida (1991), any translation will be characterized by the relative proportion (or percentage) of adaptation which it contains.

For Vinay and Darbelnet (1958, 1995) adaptation is a kind of "situational equivalence", it is a strategy that can be used when the situation mentioned in the ST cannot be found in the target culture or does not have the same importance or connotations as in the source context.

Iliescu (2017) defines translation and adaptation and argues that they are two processes very different from each other. He describes test adaptation as “a *mélange between science and practice*”. Based on Iliescu's view (2017), test adaptation includes test translation but is much more; test translation is linguistically driven (content over intent), while test adaptation is validity-driven (intent over content).

Van de Vijver (2015) differentiated between five types of adaptations: construct-driven adaptations, language-driven adaptations, culture-driven adaptations, theory-driven adaptations, and familiarity/recognisability-driven adaptations.

Equivalence is a key concept in translation assessment. Several translation scholars like Jakobson (1966); Nida (1964); Catford (1965); House (1977, 1997); Neubert (1970, 1985); Pym (1995); Koller (1995, 2011); Iliescu (2017) see equivalence as a key concept.

Jakobson (1966) spoke of ‘equivalence in difference’, Nida (1964) suggested 'different kinds of equivalence'; Catford (1965) stated that translation equivalence is essentially situational.

According to House (2015), equivalence is both a core concept in translation theory and the conceptual basis of translation quality assessment. In terms of translation quality of international assessments in education, House's theory on covert translation is very useful. House argues that a source text and its covert translation have equivalent purposes and it is possible and desirable to keep the function of the source text equivalent in the translation text and this can happen by inserting a “cultural filter” (see House, 2015, p. 68 for details). She believes that the functional equivalence comprised of ideational and interpersonal components is the basic criterion of translation quality. She also uses two other terms for these elements that are referential and non – referential.

Iliescu (2017) developed his approach to equivalence. Iliescu argues that equivalence is established via comparisons, requires a comparative logic; equivalence refers to measurement in context; equivalence should take into account the purpose of testing. He argues that equivalence

refers to a special form of validity, refers to a matter of degree, and is a property of scores, not of the instrument. Equivalence, like validity in general, is directly connected with bias and error, is a matter of generalization, and can be discussed based on the sources of validity evidence. Unlike the traditional approach considering equivalence as a single demonstration, Iliescu (2017) argue that evidence of equivalence will need to come from multiple sources, based on content, response processes, internal structure, and relations with other variables. This is a new viewpoint. Iliescu sees equivalence closely connected with biases that are associated with an error. According to him, equivalence must be seen not through the lenses of types, but the lenses of domains: language, culture, and measurement, naming linguistic equivalence, cultural equivalence, and psychometric equivalence. Analyzing the linguistic equivalence, Iliescu in consonance with Bracken and Barona (1991) and with Kristjansson et al. (2003), considers that there are three main issues generating bias in linguistic equivalence: (a) issues in lexical mapping, (b) issues in grammatical and syntactic equivalence and (c) experiential equivalence.

Bartman et al. (2018) have given a great contribution by producing tests translation and adaptation guidelines. They explain the difference between test translation from test adaptation. According to their view, translating the test is the first step towards the final adaptation of the test to the culture and psychology of the various test takers. Adaptation which sometimes is called “localization” is an extensive term and refers to bringing a test from one language and culture to another.

3. Translation quality policy in PISA, TIMSS, and PIRLS

PISA, TIMSS, and PIRLS have developed the translation policy framework that contains aims, principles, standards, and procedures. The following five basic concepts underlie the translation quality framework of these international large-scale assessments: *quality assessment*, *translatability assessment*, *quality assurance*, *quality control*, and *quality monitoring*.

Through the *quality assessment* process, PISA, TIMSS, and PIRLS have undertaken the following different measurements to find the extent to which a translated text complied with quality specifications: translatability assessment, quality assurance procedures, quality control process, and quality monitoring. All these quality procedures are implemented during instruments design, Field Trial, Main Study, and study finalization.

Based on the documents produced by PISA (2000, 2003, 2006, 2009, 2012, 2015, 2018), TIMSS (1995, 1999, 2003, 2007, 2015, 2019), PIRLS (2001, 2006, 2016), TIMSS and PIRLS (2011), like translation and adaptations guidelines, technical reports, technical standards, and research papers from the field, can be noted that the necessity of the high quality of translation and rigorous item translation review procedures are determined by a number of factors and requirements like: increased number of languages into which assessment instruments are translated; the administration of the assessment instruments in different education systems worldwide; the requirement to ensure equivalency of assessment instruments, especially equivalency in meaning and difficulty; the requirement to ensure valid testing among countries; the requirement to ensure comparability of the assessment data among countries; the requirement to keep the student mental efforts the same everywhere; the requirement to avoid heavier cognition student load; the ensure that no more working memory is needed by the student to decode and understand the given information.

The goal of quality assurance in PISA is to produce tests that yield reliable and comparable data in all countries that take part in the exam. This is accomplished by following strict procedures in achieving localization of the national versions of the testing tools employed in PISA 2018. As part of the quality control procedures, the national versions of all tests are submitted for verifications to a team of independent verifiers who see to the fact that this version is equivalent to the English and French source texts.

The TIMSS & PIRLS International Study Centre made available assessment instruments translation and adaptation guidelines. These guidelines were supposed to be followed by all the participating countries, including countries that administered the TIMSS instruments in English or used the Arabic source versions. The international source materials were translated in the language of the participating countries paying special attention to their localization and following the procedure required for quality and comparability of the national versions of the tests.

Based on the PIRLS international quality assurance program, the various national versions produced are submitted to an external review by experts in the field of linguistics and assessment. They review both the translation quality and layout of the assessment instrument. At the end of this process, the examiners provide comprehensive feedback so that the tests could be more accurate.

3.1. Standards in Translation Quality

Implementation of the translation quality in the international large-scale assessments is extremely important. From the perspective of translation, quality ensures the translation equivalency, shows that translation has been done accurately manner, and the translated text can transmit information in a similar way to the source document. This is an indicator validity and comparability of the assessment data.

The standardization of translation, adaptation, and verification procedures has been accepted as the most appropriate way that enables the fulfillment of the translation quality. Different sets of translation quality standards are developed. Translation quality standards are grouped into two main categories *metrics-oriented* that are focused on the quality metrics, that measure quality and rate it as low or high quality; and *process-oriented* that are focused on the quality translation process.

In the following, some of the standards used to ensure the translation quality of the international large-scale assessments in education are highlighted. The PISA's, TIMSS's and PIRLS's experiences show that translation quality assurance can be achieved through implementation of the translation, adaptation and verification standards and metrics as well. Based on the review of the experience of PISA, TIMSS and PIRLS we found that these programs have developed the translation quality policy and have implemented a set of translation quality standards for their purposes, based on four following international sets of standards.

ISO 17100:2015 is a special international quality standard set that defines the requirements for the translation service to meet the appropriate specifications for ensuring a qualitative translation. It is developed by the International Organization for Standardization. The ISO 17100:2015 standard also requires that the translation process must consist of at least two steps: translation and revision (<https://www.iso.org/standard/59149.html>).

The *European EN 15038* translation-services standard is a quality standard that was set by the European Committee for Standardization in a specific way for the translation industry and concentrates mainly on procedures in translation services and quality control (DIN EN 15038, 2006).

The ITC Guidelines on Translating and Adapting Tests (2005, 2017) is one of the most influential and important contributions given by the International Test Commission (ITC). It is a document that informs and directs work on adapting tests for different linguistic and cultural

contexts. ITC has become a frame of reference for obtaining high-quality test translation and adaptation. The improved second edition responded to new developments in testing technology and practice. This document contains 18 guidelines that are presented into six categories: pre-condition, test development, confirmation, administration, scoring and interpretation, and documentation. Each guideline contains an explanation and suggestions for practice. Ten of these guidelines are directly linked to the translation quality of the international large-scale assessments. These guidelines are a practical tool for conducting and evaluating the adaptation or simultaneous development of tests from the field of psychology and education, for use in various populations. Below are presented those ITC guidelines that are directly linked to the quality of the translation and implemented by PISA, TIMSS, and PIRLS.

3.2. Pre-Condition Guidelines

The PC-3 standard of this category seeks to minimize any cultural or linguistic differences that are not appropriate for use in the country concerned (*ITC, 2017*).

Based on this standard, PISA, TIMSS, and PIRLS have implemented field trial procedures to discover any linguistic differences that can impact the results of the tests.

From the category *Test development guidelines (ITC, 2017)*, there are four guidelines [TD.1 (4); TD-2 (5); TD-3 (6); TD-5 (8)] that ensure the translation quality. These guidelines are related to the development phase of the adapted test, to the appropriate translation and adaptation of all components of the test considering linguistic, psychological, and cultural differences; ensuring test adaptation suitability maximization; keeping similar meaning for intended population; and creating evidence that can support item analyses, reliability and validity in case of needed improvements. Activities recommended by the guidelines of this category are either judgmental or are based on pilot data.

The translation process is conceived to guarantee the standardization of instruments across countries. Each country is supposed to follow standardized procedures. So that to ensure the accurate translation and the international comparability of the instruments before the translated instruments are handed out to students, they undergo some strict processes of translation, verification, and review. This process was administered by the IEA Secretariat in Amsterdam for TIMSS and PIRLS and by cApStAn for PISA.

The translation, adaptation, and verification guidelines are in use by PISA, TIMSS, and PIRLS. These documents that are improved time from time help instruments developers to ensure

the equivalence and the required quality of tests, questionnaires, and other important documents for the program implementation.

Three guidelines from the *Confirmation guidelines* [C-2 (10); C-3 (11); C-4 (12)], (ITC, 2017) are linked to the translation quality. This category contains confirmation guidelines that focus on activities performed after the adapted test is considered ready for use.

The guidelines of this category emphasized the equivalence of the construct and item of the source and target versions and the locating possible sources of method bias in the adapted tests.

Administration guidelines A-1 (13), (ITC, 2017) underline the requirement for clarity of test instruction that it is a factor that can support test administration.

Guideline Doc-1 (17), (ITC, 2017) as a *documentation guideline* linked to the technical documentation of any changes presented as evidence to support equivalence, when a test is adapted for use in another population.

PISA, TIMSS, and PIRLS have implemented successfully this guideline by documenting in detail all adaptations made by the countries. They have produced special forms for the adaptation and verification processes of tests and questionnaires, like Test Adaptation Spreadsheet (TAS), Booklet Adaptation Spreadsheet (BAS), Final Optical Check (FOC), Manual Adaptation Spreadsheet (MAS), Questionnaire Adaptation Spreadsheet (QAS) used by PISA and National Adaptation Form NAF used by TIMSS and PIRLS.

Standards for Educational and psychological testing. American Educational Research Association (AERA), *American Psychological Association (APA)*, *National Council on Measurement in Education (NCM)*, 2014. As it is mentioned in the introduction of the document, this set of standards aims to support sound testing practices and to ensure a basis for quality assessment practices. Iliescu (2017) stated that the “Standards for Educational and Psychological Testing” are one of the most authoritative reference texts for the domain of testing. The fourth edition of the work, published in 2014, contains several discussions and standards that can be interpreted in the context of test adaptation”.

Standard 3.12 of the above set states that: *“When a test is translated and adapted from one language to another, test developers and/or test users are responsible for describing the methods used in establishing the adequacy of the adaptation and documenting empirical or logical evidence for the validity of test score interpretations for intended use”* (AERA, APA, NCM, 2014).

This standard requires that in the case of translation and adaptation from one language to another, the methods used for translation and adaptation should be described in detail and should be reported.

4. Translation and Adaptation Procedures

PISA, TIMSS, and PIRLS aim to create through the process of translation and adaptation instruments that are comparable to the originals in terms of equivalence of meaning and difficulty, allowing each country to adopt the materials based on national needs. This is reflected in the translation, adaptations, and verification documents prepared by three programs, like PISA 2018; Instrument Translation and Layout Verification for TIMSS 2019; Translation and Adaptation Guidelines Translation and Layout Verification for PIRLS 2016. These kinds of documents are prepared and implemented for all rounds of these programs.

To fulfil the aim of translation/adaptation, PISA (2018) asked the participating countries to meet the translation quality standards so that they would not distort the statistical results and comparisons. The purpose of the guidelines for translation and adaptation of the instruments is to ensure that the translation and adaptation of the national version present the same level of understanding and difficulty as the international version instruments (TIMSS, 2020).

“The ultimate goal of the translation and adaptation process was to create national versions of the PIRLS 2016 instruments that accommodate national languages and context while maintaining international comparability” (PIRLS, 2016).

Thus, all three programs have clearly articulated the aim of the translation, adaptation, and verification procedures that contains the common key element: keeping the quality standards, while making national adaptations to maintain international comparability.

Analyzing translation and adaptation procedures used by PISA and TIMSS we noted that PISA’s translation and adaptation guidelines are more comprehensive, whereas TIMSS quality guidelines are briefer and more general.

Comparing the translation and adaptation framework of these three international large-scale assessments can be noted that they have sets of standards, rules, and recommendations that are clearly defined, practical and useful, and time to time improved. This was remarked previously by Dept, et al. (2010).

5. Conclusion

Looking at the translation quality policy implemented by PISA, TIMSS, and PIRLS can be noted that the translation policy is influenced and is developed based on theoretical considerations of different scholars, using their thesis, considerations, methodologies, typology, and recommendation linked to linguistic quality, to linguistic equivalence, to translation, adaptation, and verification.

PISA, TIMSS, and PIRLS have advanced practical knowledge that is derived from research knowledge but based on their achievements can be concluded that these large-scale assessments have advanced translation quality theory as well. Design and implementation of unified translation; adaptation and verification standards and procedures within each program; new translation designs; development of translatability categories and verifier intervention categories; creation pools of translators and verifiers; establishment of the structures responsible for the translation quality implementation; modeling the translation, adaptation and verification documentation; numerous research papers and reports on translation quality issues; transferring translation quality procedures to national assessments are some of the indicators of the contribution given by PISA, TIMSS and PIRLS in the field of translation quality.

As translation quality is a complex issue PISA, TIMSS and PIRLS are facing some challenges that are related to cooperation among all actors involved in the translation quality, to maintain equivalence of translation avoiding bias and poor quality, to the digitalization of the translation quality process. This is a topic that requires further in-depth comparative studies.

References

- Bartman, D., Hambleton, R., Grégoire, J., & Muñiz, J. (2018). ITC Guidelines for Translating and Adapting Tests (Second Edition). *International Journal of Testing*, Volume 18, Issue 2, pp.101-134. <https://doi.org/10.1080/15305058.2017.1398166>
- Bracken, B. A., & Barona, A. (1991). State of the art procedures for translating, validating, and using psycho-educational tests in cross-cultural assessment. *School Psychology International*, 12(1-2), 119–132. <https://doi.org/10.1177/0143034391121010>
- Byrne, M. B. (2008). Testing for multigroup equivalence of a measuring instrument: A walk through the process. *Psicothema*, 20 (4), pp. 872–882.

- Catford, J. C. (1965). *A Linguistic Theory of Translation: An Essay on Applied Linguistics*, London: Oxford University Press.
- Cook, L. L., Schmitt-Cascallar, A. P., & Brown, C. (2005). Adapting Achievement and Aptitude Tests: A Review of Methodological Issues. In: Hambleton, R. K., Merenda, P. F. & Spielberger, C. D., ed., *Adapting Educational and Psychological Tests for Cross-Cultural Assessment*. Mahwah, NJ: Lawrence Erlbaum Associates, pp. 171-192.
- Dept, S., Ferrari, A., & Wäyrynen, L. (2010). Procedures in Three Multilingual Assessments: A Plea for an Integrated Translation and Adaptation Monitoring Tool. In: Harkness, J. et al., ed., *Survey Methods in Multinational, Multiregional and Multicultural Contexts*. John Wiley & Sons, Hoboken, NJ, pp.157-173. <https://doi.org/10.1002/9780470609927.ch9>
- Dept, S., Ferrari, A., & Halleux, B. (2017). Translation and cultural appropriateness of survey material in large-scale assessments. In: P. Lietz, P., Cresswell, J., Rust, K. & Adams, R., ed., *Implementation of large-scale education assessments*. Chichester, UK: Wiley, pp. 168–191. <https://doi.org/10.1002/9781118762462.ch6>
- DIN EN 15038. (2006). Translation services-Service requirements. English version of DIN EN 15038:2006-08.
- Grégoire, J., & Hambleton, R. K. (ed.). (2009). Advances in test adaptation research [Special Issue]. *International Journal of Testing*, 9(2), pp. 73-166. <https://doi.org/10.1080/15305050902880678>
- Grisay, A., de Jong, J., Gebhardt, E., Berezner, A., & Halleux-Monseur, B. (2007). Translation Equivalence across PISA Countries. *Journal of Applied Measurement*, 8 (3), pp. 249-266.
- Hambleton, K.R. & Patsula, L. (1998). Adapting Tests for Use in Multiple Languages and Cultures *Social Indicators Research*, Volume 45, pp.153–171. <https://doi.org/10.1023/A:1006941729637>
- Hambleton, R. K., Merenda, P. F., & Spielberger, C. D., ed. (2005). *Adapting educational and psychological tests for cross-cultural assessment*. Mahwah, NJ: Erlbaum. <https://doi.org/10.4324/9781410611758>
- Harkness, J. A., van de Vijver, F.J. R., & Johnson, T.P. (2003). Questionnaire design in comparative research. In: J. A. Harkness, J. F. R. van de Vijver, & P. P. Mohler, ed., *Cross-Cultural Survey Methods*, Wiley, pp. 19-34. <https://doi.org/10.4135/9781849209458.n4>

- Harkness, J.A. (2007). Improving the comparability of translations. In: Jowell, R., Roberts, C., Fitzgerald, R. & Eva, G., ed., *Measuring Attitudes Cross-Nationally: Lessons from the European Social Survey*. Sage, Los Angeles, pp. 79-93.
<https://doi.org/10.1002/9780470609927.ch3>
- Harkness, J., Edwards, B., Hansen, S. E., Miller, D., & Villar, A. (2010). Designing Questionnaires for Multipopulation Research. In: Harkness, J., et al., ed., *Survey Methods in Multinational, Multiregional and Multicultural Contexts*. John Wiley & Sons, Hoboken, NJ, pp. 33-57.
<https://doi.org/10.1002/9780470609927.ch7>
- Harkness, J., Villar, A., & Edwards, B. (2010). Translation, Adaptation, and Design. In: Harkness, J. et al., ed., *Survey Methods in Multinational, Multiregional and Multicultural Contexts*. John Wiley & Sons, Hoboken, NJ, pp. 117-140.
- House, J. (1977). (2nd ed., 1981) *A Model for Translation Quality Assessment*. Tübingen: Narr.
- House, J. (1997). *Translation Quality Assessment: A Model Revisited*. Tübingen: Narr.
- House, J. (2015). *Translation Quality assessment. Past and present*. Routledge.
<https://doi.org/10.4324/9781315752839>
- Iliescu, D. (2017). *Adapting tests in linguistic and Cultural Situations*. Cambridge University Press. <https://doi.org/10.1017/9781316273203>
- ISO 17100:2015. Translation services-Requirements for translation services.
<https://www.iso.org/standard/59149.html>
- ITC Guidelines for Translating and Adapting Tests (2nd ed., 2017). *International Journal of Testing*, <https://doi.org/10.1080/15305058.2017.1398166>
- Jakobson, R. (1966). On Linguistic Aspects of Translation. In: Brower, R., ed., *On Translation*. New York: Oxford University Press.
- Kelly, L. D., Malak, B. (2003). Translating the PIRLS Reading Assessment and Questionnaires. In: Martin, O. M., Mullis, S.I., & Kennedy, M.A., ed., *PIRLS 2001 Technical Report*. International Study Center, Lynch School of Education, Boston College pp. 44.
- Koller, W. (1995). The Concept of Equivalence and the Object of Translation Studies. *Target 7*: pp. 191–222. <https://doi.org/10.1075/target.7.2.02kol>

- Koller, W. (2011) (11th ed.) *Einführung in die Übersetzungswissenschaft*. Tübingen: Francke.
- Kristjansson, E. A., Desrochers, A., & Zumbo, B. (2003). Translating and adapting measurement instruments for cross-linguistic and cross-cultural research: A guide for practitioners. *Canadian Journal of Nursing Research*, 35 (2), pp. 127–142.
- Maxwell, B. (1996). Translation and Cultural Adaptation of the Survey Instruments. In: Martin, O.M. & Kelly, L. D., ed., *Third International Mathematics and Science Study (TIMSS) Technical Report, Volume I: Design and Development*. IEA, pp.8-1-9-1.
- Neubert, A. (1970). Elemente einer allgemeinen Theorie der Translation. In: *Actes du Xe Congrès International des Linguistes*, Bucharest, 1967, pp. 451–56.
- Neubert, Albrecht (1985) *Text and Translation*. Leipzig, Enzyklopädie.
- Nida, E. (1964). *Towards a Science of Translating*. Leiden: E.J. Brill.
<https://doi.org/10.1163/9789004495746>
- Nida, E. 1991. Theories of translation. *TTR: Traduction, terminologie, redaction*, vol.4, no1, 1991, pp. 19-32. <https://doi.org/10.7202/037079ar>
- Nida, E. and Taber, C.R. (1969). *The Theory and Practice of Translation*. Leiden: E.J. Brill.
- Nord, C. (1997). A functional typology of translations. In: Trosborg, A., ed., *Text typology and translation*, John Benjamins, pp. 43–66. <https://doi.org/10.1075/btl.26.05nor>
- PIRLS 2016. Ebbs, A. & Wry, E. *Translation and Layout Verification for PIRLS 2016*. TIMSS & PIRLS International Study Center, pp.7.1.
- PISA 2018. Translation and adaptation guidelines. Produced by cApStAn and Béatrice Halleux under the supervision of ETS, Core A and Core D Contractors, pp.1.
- Pym, A. (2010). *Exploring Translation Theories*. London and New York: Routledge, p. 324.
<https://doi.org/10.4324/9780203869291>
- Reiss, K. (2000). *Translation criticism: The potentials and limitations. Categories and criteria for translation quality assessment*. Routledge.
- Rios, A. J., & Sireci, G.S. (2014). Guidelines Versus Practices in Cross-Lingual Assessment: A Disconcerting Disconnect. *International Journal of Testing*, Volume 14, Issue 4, pp.289-312 <https://doi.org/10.1080/15305058.2014.924006>

- Sireci, S. G. (2005). The most frequently unasked questions about testing. In: Phelps, P.R., ed., *Defending standardized testing*, Lawrence Erlbaum Associates Publishers, pp. 111–121.
- Standards for Educational and Psychological Testing*. (2014). Developed jointly by American Educational Research Association, American Psychological Association, National Council on Measurement in Education, pp. 68.
- TIMSS. (2020). Instrument translation and Layout Verification for TIMSS 2019. In: Martin, O.M., Davier von M., & Mullis, S.V. I., ed., *Methods, and procedures: TIMSS 2019 Technical report*. TIMSS & PIRLS International Study Center, pp. 5.3.
- van de Vijver, F. J. R. (2015). Methodological aspects of cross-cultural research. In: M. Gelfand, Y. Hong., & C. Y. Chiu., ed., *Handbook of advances in culture & psychology* (Vol. 5, pp. 101-160). New York, NY: Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780190218966.003.0003>
- van de Vijver, F. J. R., & Leung, K. (1997). *Methods and data analysis for cross-cultural research*. Thousand Oaks, CA: Sage.
- Vinay, J. P., & Darbelnet, J. (1995). *Comparative Stylistics of French and English: A Methodology for Translation. North America, Amsterdam, and Philadelphia*. John Benjamins
<https://doi.org/10.1075/btl.11>
- Vinay, J.P., & Darbelnet, J. (1958). *Stylistique Comparée du Français et de l'Anglais: Méthode de Traduction*. Paris: Didier. (Transl. and ed. by Sager, J.C. and Hamel, M.J. (1995) as *Comparative Stylistics of French and English: A Methodology for Translation*. John Benjamins.
- Williams, M. (2004). *Translation quality assessment: An Argument-centered approach*. Ottawa: University of Ottawa Press.
- Williams, M. (2009). Translation quality assessment. *Mutatis Mutandis* 2 (1), pp.1-21.