EUROPEAN ACADEMIC RESEARCH
Vol. IX, Issue 8/ November 2021

Impact Factor: 3.4546 (UIF)
DRJI Value: 5.9 (B+)

ISSN 2286-4822
www.euacademic.org

Translation verification in international large-scale assessments in education

LUELA LIÇI, MSc
Ph.D. candidate, University of Tirana
Albania

Abstract

Translation verification is a key part of the PISA and TIMSS international quality assurance programs. This paper aims to analyze the translation verification frameworks used by PISA and TIMSS focusing on aims, processes, quality metrics models, achievements, and challenges. Based on this research, while PISA and TIMSS, have similar translation verification goals, they have differences in terms of the verification process and quality metric model. PISA and TIMSS intend to provide feedback to understand the quality and comparability of the translated/adapted instruments by participating countries; to improve the accuracy and comparability of the instruments to maintain the same meaning and level of difficulty of the target language as the international version; to produce high-quality translations that are internationally comparable. The research found as well that PISA and TIMSS are using different translation verification methodologies. TIMSS verification methodology is based on severity codes, while PISA's verification process is based on Multidimensional Quality Metrics. This paper concludes that despite the differences, translation verification has been successful and has improved the quality of the translation of international assessment instruments in education. The translation verification guidelines, data from the technical reports of these international assessment studies, and verification results are used as the sources for this paper.

Keywords: translation, translation equivalence, verification procedure, severity code, multidimensional quality metrics

INTRODUCTION

During the past 20 years, the interest in international large-scale studies on assessment of the academic achievements, the number of

participating countries, and the number of languages in which the assessment instruments have been translated have increased. In this context, the team of the Organisation for Economic Co-operation and Development (OECD) that is responsible for the Program for International Student Assessment (PISA) and the team of the International Association for the Evaluation of Educational Achievement (IEA) that is responsible for the Trends in International Mathematics and Science Study (TIMSS) have produced technical standards and guidelines to ensure the quality of their instruments in order to enable comparative studies on educational achievements. In the framework of the translation quality policy, translation verification is an important part of the translation quality policy, ensuring the translation equivalence of the instruments of the international PISA and TIMSS programs.

Grisay et al. (2007) noted that in the context of an increased number of participating countries in PISA and TIMSS, ensuring the linguistic and cultural equivalence has become the main challenge.

Martin et al. (1999) and Korsnakova et al. (2020) have recognized that the quality of IEA studies depends on the accuracy of the translation instruments, like tests, questionnaires, and manuals from the source to the target language.

Martin et al. (1999) pointed out that the verification of the translated instruments into the languages of the participating countries is perceived as a measure ensuring the comparability of the collected data from the achievement performance from different countries and cultures, by not changing the meaning and the difficulty level of the items from the international version.

Korsnakova et al. (2020) mentioned that ensuring linguistic equivalence of national versions is very important after ensuring the validity and reliability of the international source version.

IEA studies are guided by several standards documents. *Standards for Education Data Collection and Reporting* (SEDCAR) (1991), produced for the United States National Centre for Educational Statistics (NCES) by Westat contain, among others, translation and verification translation standards. The guiding translation standard is set as follows:

"When translating test items or modifying them for cultural adaptation, the following must remain the same as the international version: the meaning of the question; the reading level of the text; the

difficulty of the item; and the likelihood of another possible correct answer for the test item" (Martin et al.1999, p.43).

The PISA 2021 Guidelines on Translation and Adaptation make it clear that the translation equivalence of national versions in accordance with the international version is required to be provided in order to avoid bias and to enable international comparison.

Verification of the translation equivalence is required even if the translation can be adequate, the verification brings improvement, suggests another wording that better guarantees equivalence and meets the international standards of translation quality.

In the case of the PISA and TIMSS programs, translation verification is a standardized procedure. The clear quality verification procedures are defined, translation quality standards are defined, metric systems are selected and guidelines are prepared. A variety of ways are used to perform translation verification, like back translation model, double translation model, translation verification by country editing, verification by a language quality control company, verification of the quality by independent translators, international verification.

METHODOLOGY

Product-oriented research methodology in translation is used for the purpose of this research. A qualitative analysis has been performed. The translation verification guidelines produced by PISA and TIMSS; data from the technical reports of these international assessment studies, research results presented in different papers and books are used as the sources for this paper. The issue is studied in detail, products are analysed and interpretations are presented.

The purpose of translation verification

Translation quality of PISA and TIMSS is developed using the equivalence-based approach, focusing mainly on the equivalence of meaning and equivalence of difficulty.

According to the translation verification guidelines of PISA 2021 and TIMSS 2019, the aim of the translation verification is to guarantee translation equivalence of assessment instruments to maintain the same meaning and reading level of difficulty of the target language as the international version and to ensure that

students in different countries use the assessment instruments in the same way and so the assessment provides reliable and fully comparable information. While there are the following purposes of the translation verification:

- to provide feedback to understand the quality and comparability of the translated/adapted instruments by participating countries;
- to improve the accuracy and comparability of the instruments to maintain the same meaning and level of difficulty of the target language as the international version;
- to produce high-quality translations that are internationally comparable.

The final goal of the translation and verification process of the PISA and TIMSS is to produce the national versions instruments that render national context and language instruction while keeping international comparability.

According to Korsnakova et al. (2020), IEA procedures aim to maximize the comparability of the data and to ensure that the biggest errors or misunderstandings are avoided.

Based on the PISA 2021 Translation and Adaptation Guidelines, the verification process of the equivalence is focused on: fulfillment of the standards of keeping the same difficulty (not reducing or increasing) of the text comprehension, graphics, tables, and stimulus of the tests; on avoiding ambiguities in the questionnaires; on keeping of the manuals adaptations to the local context at the level that does not impact the change of the collected data.

Practically, PISA and TIMSS intend to implement the agreed standards and to check if the specifications have been fulfilled. For this purpose, there are produced translation verification guidelines and is institutionalized the translation verification system and process.

Translation verification process

In order to ensure the linguistic and cultural equivalence as an important indicator of the success of the international comparative studies in education, rigorous procedures of translation and adaptation verification of the assessment instruments were implemented by PISA and TIMSS.

Several documents that describe recommended translation, adaptation, and verification procedures are produced and are in use by participating countries.

Korsnakova et al. (2020) inform that at first, TIMSS 1995 produced translation procedures based on Hambleton's (1993) recommendations. To verify the translations, TIMSS 1995 relied on multiple forward translations, translation reviews by bilingual judges (translation verifiers.

Analyzing the translation verification process implemented by IEA, Korsnakova et al. (2020) presented a list of the instrument preparation that contains nine stages and responsibilities of the international study center, national centers national translators, national reviewers, and international verifiers. From their research, Korsnakova et al. found that international verification consists of three steps: adaptation verification, translation verification, and layout verification, and the order of the verification steps has changed over the years. They remarked that IEA studies implemented a decentralized approach for translation and adaptation of the assessment instruments at the national level, while a centralized approach was adopted for international verification.

At the beginning, the PISA Translation and Adaptation Guidelines were composed based on the literature on international test adaptation (Hambleton, 1993, 1994; Hambleton and Patsula, 1998; Hambleton and Jong, 2003; Jeanrie and Bertrand, 1999,), on IEA similar guidelines developed for TIMSS (O'Connor and Malak, 2000;), and later the guidelines are improved using the PISA experience as well (PISA, 2021).

Table 1 presents a summary of the stages of the translation verification process implemented by PISA 2018 and TIMSS 2019. Research on the translation verification procedure of PISA and TIMSS found that both programs have the same aim and purposes for the translation verification process, while the ranking of the stages differs slightly. TIMSS groups the translation verification procedures using two major sections: translation verification and layout verification. The order in which these verification steps are conducted has changed over the years (for more information see Korsnakova et al., 2020, pp. 98-99).

Table 1. Translation verification process

PISA 2018	TIMSS 2019
Translation Double translation either from both English and French source versions.	*Translation verification* During translation verification, verifiers provide detailed feedback and ask for improvements. After the National Research Coordinators (NRCs) implement required changes, countries submitted their national instruments to the TIMSS & PIRLS International Study Centre for layout verification.
Reconciliation/Adaptation VF Report on issues to be considered in the national materials.	*Review* The review process is focused on the evaluation of the readability and accuracy of the translated material towards the target population.
Verification Preparation Preliminary checks. In the case of any major problem is remarked then materials can be returned for review.	*Translation and Adaptation of the Achievement Instruments* During this stage attention is paid to terms that convey the same meaning, to the style of text, ensuring that the meaning and difficulty of the item remained the same as to the international version question. Consistency of adaptations and translations from item to item must be maintained.
Verification During this step, the verifier makes all changes and interventions to the national version.	*Translation and Adaptation of the Context Questionnaires* Participating countries are required to use terms that are appropriate for the national context and education system.
Verification review cApStAn staff reviews the verifier feedback.	*International Translation Verification* After the instruments are translated and adapted, they must be submitted to IEA Amsterdam for translation verification.
Referee review Translation referee of the country reviews the comment and verifier feedback.	*The Translation Verification Process* • Checking the accuracy, linguistic correctness, and comparability of the translation and adaptations of the items and questionnaires. • Documenting any deviations between the national and international versions. • Suggesting a translation/adaptation alternative to improve the accuracy and comparability of the national instruments. • Verifiers provide feedback on the quality of the translated and adapted texts.
Post-verification review National Centre (NC) based on the verifier and referee's feedback finalizes the national version.	*Translation Verification of the Trend Assessment Blocks* • Ensuring that the trend items had not changed. • Recording any discrepancies found in the trend items in the NAF, eTIMSS Online Translation System, or Bridge Verification Form. • NRCs were required to carefully review all discrepancies and discuss any proposed changes with the TIMSS & PIRLS International Study Centre.
Test developer review Test developer's review on item-specific comments in the TAS, made by NC, verifier or Referee, and related follow-up.	*Review of International Translation Verification Feedback* At this stage after completion of international translation verification, the NRCs are responsible for responding to the translation verifiers' feedback by either accepting, modifying or rejected suggested changes to the adapted and/or translated text.
Layout adaptation Layout review by Core 3 (ETS) to fix any residual layout issue documented in the TAS.	*Layout Verification* Following translation verification, all national instruments are required to undergo layout verification by the TIMSS & PIRLS International Study Centre. Layout verification is the final external review.
Test developer sign-off Test developer's sign-off, once the agreement was found on item-related issues. Files were then sent back to cApStAn for the Final Check.	*Layout Verification of Achievement Materials* The primary goal of layout verification of achievement materials is to ensure that students in different countries experience the assessment instruments in the same way.
Final check preparation Preliminary checks on the materials submitted by the NC after verification and preparation of the materials for FC.	*Layout Verification of Context Questionnaires* The context questionnaires are checked against the international versions to identify any potential layout issues as well as to ensure the international comparability of the questionnaire data. During layout verification of questionnaires, the verifiers take into consideration any national adaptations documented by the NRCs.
Final check Correctness of any post-verification changes made by the	*Layout Verification of Trend Materials and Bridge Booklets* During layout verification of trend materials, the verifiers ensure that the layout structure and adaptations in the national TIMSS 2019 instruments are

NC and correct implementation of the latest errata.	consistent with countries' trend versions.
Final check review Review of the verifier feedback by cApStAn staff.	*Review of Final Instruments* At this final stage are foreseen necessary adjustments to the materials and responding to all the feedback from the layout verifiers. NRCs were required to submit their materials to the TIMSS & PIRLS International Study Center for a final review. Once, the TIMSS & PIRLS International Study Center confirmed the materials were finalized, the country is permitted to begin printing the paper-based instruments.
Final review NC review of verifier's comments arising from Final Check and if the further layout or text changes were needed, NC requested to implement them.	

The research found that verification is a focused process for both international studies. The parts of any test booklet or a questionnaire that are critical towards the equivalence are selected for verification.

Linguistic quality control is a part of the translation verification process in PISA and TIMSS. It includes verification by linguistics, documentation of issues, monitoring of correction, and preparation of the final report of the verification process. Dual verification is another feature of the process. A linguist and a subject matter expert work together and a project manager combines their feedback into a workable report.

The verification process in both cases takes place in two stages. The first stage is the preparatory one and it is performed before the translation starts. The components of this stage are preparation of the translation and adaptation guidelines, preparation of the verification forms, and training of verifiers. In the post-translation phase, the verification of the translated materials is performed.

Since differences in the layout can also affect the international comparability of the data, PISA and TIMSS conduct the verification of the national instrument layout. During layout verification, the national instruments are compared to the international instruments, and any discrepancies are documented and corrected. The goal of layout verification is to ensure minimal deviations in the comparability of the layout of national instruments.

Research proves that PISA and TIMSS have established a translation verification system composed of entities and bodies, with clear tasks and responsibilities, well-structured and ruled by standards. The verification process of PISA and TIMSS is an institutionalized and structured activity. As presented in Figure 1, the

verification process is implemented by cooperation of specialized entities like the Australian Council for Educational Research (ACER) and International Association for the Evaluation of Educational Achievement (IEA); quality control centers like cApStAn that is responsible for the translation verification of all national versions of PISA survey instruments since 2000 and of TIMSS 2007 for a number of national versions and of TIMSS 2011, 2015, and 2019 of all national versions; TIMSS & PIRLS International Study Center; national centers; translators, verifiers, reviewers; and by using standardized documents like Test Adaptation Spreadsheet (TAS), Questionnaire Adaptation Spreadsheet (QAC), Final Optical Check (FOC), National Adaptation form (NAF) for that recording all linguistic deviations, giving feedback and recommendations.

Figure 1. PISA translation verification system

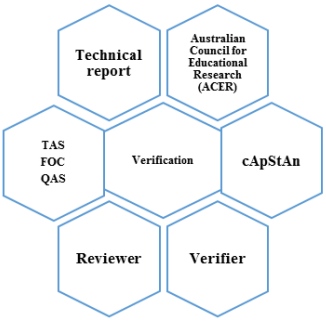
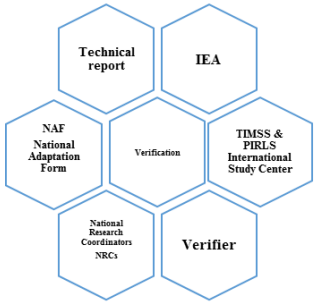


Figure 2. TIMSS translation verification system



Linguistic quality control

The main goal of the linguistic quality control (LQC) of PISA and TIMSS is to evaluate the accuracy of the translation, the justification for and adequacy of any cultural adaptations, and the comparability of the layout of the assessment instruments. The practical task of translation verifiers is to ensure accuracy and the highest quality of the localized content.

Taking into consideration this goal, PISA and TIMSS have selected the error metric methodology, practically aiming to identify errors ranging from simple, such as spelling errors, to those which are

harder, such as mistranslation, inconsistent translations, the literal translation of idioms, or colloquial expression, incorrect translation or missing terminology.

Research has found that these international assessment programs have some different experiences towards using error metrics methodologies. Referring to the TIMSS experience can be remarked that during many years TIMSS implemented the error metric methodology based on the "type of codes" and the "severity codes", while PISA's verification process is based on the Multidimensional Quality Metrics (MQM) model. But, from 2007 partially and from 2011 mainly, TIMSS is in cooperation with cApStAn and being under the influence of this linguistic quality control company is using a mixed error metric methodology.

Table 2. Verifier intervention categories

PISA 2021	TIMSS 2019
OK	Adaptation
Added information	Comprehension
Missing information	Erratum/Update missed
Matches and patterns	Formatting
Inconsistency	Grammar
Adaptation issue	Inconsistency
Register / wording issue	Missing translation
Grammar / syntax issue	Punctuation, symbols, and spelling
Mistranslation	Redundancy
Guideline not followed	Translation issues
Left in a source language	
Minor linguistic defect	
Erratum / update missed	
Layout / format issue	

Table 3. TIMSS 2019 Codes used in verification feedback

Code	Description of the code
Code 1	*Major change or error:* These changes could affect the results. Examples include the omission or addition of a question or answer option; the incorrect translation that changes the meaning or difficulty of the item or question; and incorrect order of questions or answer options in a multiple-choice question. If in any doubt, verifiers are instructed to use CODE 1? so that the error can be referred to the TIMSS & PIRLS International Study Center for further consultation
Code 2	*Minor change or error:* Indicates a minor change or error, such as a spelling or grammar error that does not affect comprehension
Code 3	*Suggestions for alternative:* Is used when the verifier considers the translation adequate but suggests an alternative wording. The translation may be adequate, but the verifier suggests a different wording.
Code 4	*Acceptable changes:* Indicates an adaptation that is acceptable and appropriate. For example, a reference to winter is changed from January to July for a country in the Southern Hemisphere Used to identify and document that national conventions have been properly documented and implemented.

Despite the differences, translation verification has been successful and has improved the quality of the translation of international assessment instruments in education which has enabled comparable test data and valid assessment results.

The technical reports of PISA 2000, 2003, 2006, 2009, 2012, 2015, and TIMSS 2015 mentioned that the verifiers identified errors that would have seriously affected the functioning of specific items – mistranslations, omissions, loan translations or awkward expressions, incorrect terminology, poor rendering of graphics or layout, errors in numerical data, grammar, and spelling errors.

The procedures of verifying translations in the TIMSS 2015, 2019 were effective. In most cases they confirmed that national centers had produced high-quality translations; in other cases, they alerted the centers to flaws in translations in time to make changes.

CONCLUSIONS

The translation verification process of PISA and TIMSS resulted to be an essential mechanism for ensuring quality.

The research found that translation verification is an institutionalized activity. PISA and TIMSS have developed and implemented a translation verification system. Linguistic quality control is a subset of verification. Translation verification is the responsibility of the linguistic quality control centers that are cooperating with PISA and TIMSS.

Translation verification is a standardized, structured, and documented activity. Unified translation, adaptation, and verification procedures within each program PISA and TIMSS are implemented. The test verification process of PISA and TIMSS is being fully documented. Since PISA 2006, the verification process is more structured.

The research found similarities and differences in the translation quality policy, framework, and guidelines of the translation verification process implemented by PISA and TIMSS. While PISA and TIMSS, have similar and common translation verification aims and purposes, they have some differences in terms of the verification process and quality metrics model. PISA and TIMSS have implemented innovation in the field of translation verification.

cApStAn developed for PISA 2006 a set of verifier intervention categories that is improved later.

Computer-based assessments implemented by PISA, and TIMSS, are supported by e-Assessment systems to increase operational efficiency in translation verification as well.

Despite achievements, the PISA and TIMSS are facing some challenges related to translation verification: a need for better cooperation among all actors involved in the translation verification procedure; hiring better quality translators; special training of translators and verifiers on error metric methodology; digitalization of the translation verification process.

Computer-based assessments implemented by PISA and TIMSS are supported by e-Assessment systems to increase operational efficiency in translation verification as well.

Translation verification as a part of international large-scale studies should be researched in more depth in the future as it is believed that experience will be generalized and recommendations will improve the translation process.

REFERENCES

1. Ebbs, D. et al. (2019). Instrument translation and layout verification for TIMSS 2019. IEA.
2. Grisay, A., de Jong, J., Gebhardt, E., Berezner, A., & Halleux-Monseur, B. (2007) Translation Equivalence across PISA Countries. *Journal of Applied Measurement*, 8 (3), 249-266)
3. Hambleton, K.R. & Jong, D.H. Advances in Translating and Adapting Educational and Psychological Tests. *Language Testing* · April 2003 DOI: 10.1191/0265532203lt247xx
4. Hambleton, K.R. & Patsula, L. (1998). Adapting Tests for Use in Multiple Languages and Cultures *Social Indicators Research* volume 45, pages153–171)
5. Hambleton, R. K. (1994). Guidelines for adapting educational and psychological tests: A progress report. *European Journal of Psychological Assessment, 10*(3), 229–244.
6. Hambleton, Ronald K.(1993). Translating Achievement Tests for Use in Cross-National Studies. International Association for the Evaluation of Educational Achievement, New York, NY.; National Centre for Education Statistics (ED), Washington, DC.
7. Jeanrie, Ch. & Bertrand, R. Translating Tests with the International Test Commission's Guidelines: Keeping Validity in Mind. *European Journal of Psychological Assessment*, September 1999, 15(3):277-283 DOI:10.1027//1015-5759.15.3.277

8. Korsnakova, P., Dept. S. & Ebbs, D. (2020). Translation: The Preparation of National
Language Versions of Assessment Instruments. In: Wagenaker, H. (ed)., Reliability and Validity of International Large-Scale Assessment. IEA Research for Education. Pp. 85-110. https://doi.org/10.1007/978-3-030-53081-5_6
9. Martin, M. O., Rust, K., & Adams, R. J. (Eds.). (1999). *Technical standards for IEA studies*. Amsterdam, the Netherlands: IEA, p.43. Retrieved from https://www.iea.nl/publications/iea-reference/ technical-standards-iea-studies.
10. O'Connor, K.M., & Malak, B. (2000). Translation and cultural adaptation of the TIMSS instruments. In M.O. Martin, K.D. Gregory, & S.E. Stemler (Eds.). TIMSS 1999 technical report (pp. 89-100). Chestnut Hill, MA: Boston College.
11. PISA 2000 Technical Report, edited by Adams, R. & Wu, M. OECD. Published in 2002, pp.63-67.
12. PISA 2003 Technical Report.OECD.Published in 2005, p.74.
13. PISA 2006 Technical Report.OECD. Published in 2009. Pp.96-104.
14. PISA 2009 Technical Report.OECD. Published in 2012, p.96.
15. PISA 2012 Technical Report.OECD. Published in 2014, p.104.
16. PISA 2015 Technical Report.OECD. Published in 2017, p.99.
17. PISA 2018 Translation and Adaptation Guidelines. (2018) Produced by cApStAn and Béatrice Halleux under the supervision of ETS and Core A Contractors.
18. PISA 2021 Translation and Adaptation Guidelines. (2019). Produced by cApStAn and Béatrice Halleux under the supervision of ETS and Core A and Core D Contractors Contractors.
19. TIMSS. (2015). Translation and Translation Verification for TIMSS 2015. In: Martin, O.M., Mullis, S.V. I., Hooper, M. ed., *Methods and procedure in: TIMSS 2015.IEA*, p.7.16.
20. TIMSS. (2019). 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, p.5.23.