



Chapter 4

Translation and Cultural Adaptation of the TIMSS 2003 Instruments

Steven J. Chrostowski and Barbara Malak

4.1 Overview

The TIMSS 2003 data collection instruments (achievement tests and background questionnaires) were developed and prepared in English by the TIMSS & PIRLS International Study Center (ISC) at Boston College, with contribution from the National Research Coordinators (NRCs) of participating countries. The assessment instruments were subsequently translated by the participating countries into their local languages of instruction, 34 in total. Of the 49 countries and four Benchmarking participants in the TIMSS 2003 data collection, 17 collected data in two languages and one in three languages. The most common languages of testing were English (18 countries) and Arabic (10 countries).

The translation process was designed to ensure standardization of instruments across countries. Each country was expected to follow procedures established by the ISC for translating the test instruments into the national language and cultural context. These guidelines were provided to all NRCs in the *TIMSS 2003 Survey Operations Manual* (TIMSS, 2002a), and were further elaborated and discussed at relevant NRC meetings.

Before the translated instruments were administered to students, they went through a rigorous process of translation verification and review to ensure that they were translated accurately and were internationally comparable. This process was managed by the IEA Secretariat in Amsterdam. As a critical part of the translation verification process, the translated instruments for each country were checked by independent verifiers against the TIMSS 2003 international version to assess the comparability of translation. Verifiers

reviewed the translated instruments and documented any deviations from the international version. National Research Coordinators received a Translation Verification Report that identified corrections or improvements considered necessary by the verifiers. When all necessary corrections had been implemented by NRCs, the International Study Center reviewed the revised instruments, suggested additional improvements, and gave final approval to the countries to print and administer the materials.

Translation verification was conducted both for the TIMSS 2003 field test and the main data collection.¹ For the achievement tests, the bulk of the translation effort took place prior to the field test, as there were few changes to the test items selected from the field test for use in the main data collection. The background questionnaires, however, were substantially revised after the field test and therefore required a second major translation effort. For the 44 participants in the field test, verification was conducted at both stages of the study. This allowed these countries to practice the translation procedures prior to the main data collection. It also gave them an additional opportunity to check the translations of items used in both the field test and main data collection.

All countries that participated in TIMSS 2003 submitted their most important instruments for translation verification. However, some countries did not submit for verification instruments in languages which were administered to a very small proportion of the sample. Such countries, however, used instruments that were translated and verified for another country (for example, Egypt used Lebanon's French and English instruments in a few schools).

4.2 Translation of Instruments

The TIMSS 2003 survey translation guidelines called for two independent translations of each test instrument from English to the target language. A translation review team then reviewed and compared the two translations to arrive at a final version of the translated instruments.

The prescribed translation procedure at the National Research Centers included the following steps:

1. Identify the target language(s), i.e. the language(s) of instruction.
2. Identify translators for two independent translations.
3. Translate instruments and adapt as necessary.
4. Confer and reconcile the two independent translations.
5. Document all cultural adaptations.

¹ The TIMSS 2003 field test was conducted during April-June 2002, and the main data collection was conducted during September-November 2002 for southern hemisphere countries and February-July 2003 for northern hemisphere countries.

In practice, because of scarcity of resources and/or time allotted for translation, several countries used only one person to translate the instruments, often the NRC, who generally was the person most competent for this task.

4.2.1 Instruments To Be Translated

Each country had to translate the following materials into the language of instruction at each grade:

- 14 blocks of mathematics achievement items and 14 blocks of science achievement items (see next section);
- the student directions for the assessment;
- the background questionnaires – Student Questionnaire, Teacher Questionnaire, and School Questionnaire;²
- the School Coordinator Manual;
- the Test Administrator Manual, including the Test Administration Form; and
- the Scoring Guides for the Constructed-Response Items.

Countries testing in English did not have to translate the instruments, but were required to adapt the American-English of the originals to the vernacular, and make whatever adaptations were necessary for cultural reasons. The mathematics and science tests and the background questionnaires underwent the translation verification process, whereas the manuals and scoring guides did not. The International Study Center provided each country with electronic files containing all of the material to be translated to facilitate the translation.

4.2.2 Identification of the Target Language

Each NRC identified the language or languages to be used for testing (see Exhibit 4.1) and the geographical or political areas associated with them. If a single translation was prepared within a country, translators needed to ensure that the translation was acceptable to all of the dialects of the language in which the assessment was to be administered. Professionals from these dialects were to be involved in adapting the instruments and testing materials.

² At the eighth grade only, there are different versions of the student questionnaire for countries that teach science as a single general/integrated subject and for countries that teach science as separate subjects at the eighth grade, and there are separate versions of the teacher questionnaire for mathematics and science teachers.

Exhibit 4.1 TIMSS 2003 Translation Verification

Country	Grade 8	Grade 4	Language(s) of Test	Materials Verified
Argentina	√		Spanish	Adapted Chilean version of test booklets and questionnaires
Armenia	√	√	Armenian	Translated test booklets and questionnaires
Australia	√	√	English	Adapted international English version of full set of instruments
Bahrain	√		Arabic, English	Adapted Egyptian Arabic version of booklets and questionnaires
Belgium (Flemish)	√	√	Dutch	Translated full set of instruments
Botswana	√		English	Translated full set of instruments
Bulgaria	√		Bulgarian	Translated full set of instruments
Chile	√		Spanish	Translated full set of instruments
Chinese Taipei	√	√	Chinese	Translated full set of instruments
Cyprus	√	√	Greek	Translated full set of instruments
Egypt	√		Arabic, English, French	Translated Arabic version of test booklets and questionnaires
England	√	√	English	Adapted international English version of test items and questionnaires
Estonia	√		Estonian, Russian	Translated full set of instruments in both languages
Ghana	√		English	Adapted international English version of full set of instruments
Hong Kong, SAR	√	√	Chinese, English (grade 8 only)	Translated full set of instruments in Chinese and adapted international English version of questionnaires
Hungary	√	√	Hungarian	Translated full set of instruments
Indonesia	√		Indonesian	Translated full set of instruments
Iran, Islamic Rep. of	√	√	Farsi	Translated full set of instruments
Israel	√		Hebrew, Arabic	Translated full set of instruments in Hebrew, translated test blocks, test booklets, and student questionnaire in Arabic (teacher and school questionnaires not administered in Arabic)
Italy	√	√	Italian	Translated full set of instruments
Japan	√	√	Japanese	Translated full set of instruments
Jordan	√		Arabic	Translated full set of instruments
Korea, Rep. of	√		Korean	Translated full set of instruments
Latvia	√	√	Latvian, Russian	Translated full set of instruments in Latvian
Lebanon	√		French, English	Translated French and adapted international English versions of test booklets and questionnaires
Lithuania	√	√	Lithuanian	Translated full set of instruments
Macedonia, Rep. of	√		Macedonian, Albanian	Translated full set of instruments in both languages
Malaysia	√		Malay	Translated full set of instruments
Moldova, Rep. of	√	√	Moldavian, Russian	Adapted Romanian and Russian versions of test booklets and questionnaires

Exhibit 4.1 TIMSS 2003 Translation Verification (...Continued)

Country	Grade 8	Grade 4	Language(s) of Test	Materials Verified
Morocco	√	√	Arabic	Translated test booklets and questionnaires
Netherlands	√	√	Dutch	Translated full set of instruments
New Zealand	√	√	English, Maori (grade 4 only)	Adapted international English version of full set of instruments, translated test blocks and student questionnaire in Maori (teacher and school questionnaires not administered in Maori)
Norway	√	√	Bokmål, Nynorsk	Translated full set of instruments in both languages
Palestinian Nat'l Auth.	√		Arabic, English	Adapted Jordanian Arabic version of test blocks and questionnaires
Philippines	√	√	English	Adapted international English version of full set of instruments
Romania	√		Romanian, Hungarian	Translated full set of instruments in Romanian
Russian Federation	√	√	Russian	Translated full set of instruments
Saudi Arabia	√		Arabic	Adapted Egyptian version of test booklets and questionnaires
Scotland	√	√	English	Adapted international English version of test items and questionnaires (tests same version as England)
Serbia	√		Serb	Translated full set of instruments
Singapore	√	√	English	Adapted international English version of full set of instruments
Slovak Republic	√		Slovak, Hungarian	Translated full set of instruments in both languages
Slovenia	√	√	Slovene	Translated full set of instruments
South Africa	√		English, Afrikaans	Adapted international English and translated Afrikaans versions of full sets of instruments
Sweden	√		Swedish	Translated full set of instruments
Syrian Arab Republic	√		Arabic	Adapted Egyptian version of test booklets
Tunisia	√	√	Arabic	Translated test booklets and questionnaires
United States	√	√	English	Adapted international English version of test items and questionnaires
Yemen		√	Arabic	Adapted Egyptian version of test booklets and questionnaires
Benchmarking Participants				
Basque Country, Spain	√		Basque, Castilian	Translated full set of instruments in both languages
Indiana State, US	√	√	English	Adapted international English version of test items and questionnaires (same version as United States)
Ontario Province, Can.	√	√	English, French	Adapted international English and translated French versions of full sets of instruments
Quebec Province, Can.	√	√	English, French	Adapted international English and translated French versions of full sets of instruments

Note: Full set of instruments consists of test blocks, test booklets, background questionnaires, and trend items if applicable.

4.2.3 Identification of Translators for Two Independent Translations

Translators were expected to have an excellent knowledge of both English and the target language and experience in the subject matter. For the achievement tests, the translation procedure required four translators for each target language, two with expertise in mathematics education and two in science education. Where subject matter experts were not available to act as translators, the translators were expected to work closely with subject matter specialists to ensure that the content and difficulty of the items did not change as a result of the translation. If a country could not employ all the required translators, the NRC played a major role in translating and/or verifying the translation of the instruments.

Translators of general text materials (student, teacher, and school questionnaires, and procedural manuals) did not need to be subject-matter specialists, so only two translators were necessary for these documents.

4.2.4 Translation and Cultural Adaptation of Instruments

Translators were provided with guidelines and procedures to follow in translating the data collection instruments and adapting them to their national cultural context. The guidelines were designed to yield translations that were as close as possible to the international (English) version of the survey instruments, while allowing for cultural adaptations where necessary. Translators were cautioned not to change the meaning or the difficulty level of an achievement item during the translation process. The primary concern was to convey the same meaning and style of the items as closely as possible to the international version.

The translators' tasks included:

- identifying and minimizing cultural differences;
- finding equivalent words and phrases;
- ensuring that the reading level was the same in the target language as in the original international version;
- ensuring that the essential meaning of the text did not change;
- ensuring that the difficulty level of achievement items did not change; and
- making changes in the instrument layout required due to translation.

As described in Chapter 2, the TIMSS 2003 assessment uses a matrix-sampling technique that involves dividing the entire item pool into a set of unique item blocks, distributing these blocks across a set of test booklets, and rotating the booklets among the students. To facilitate the creation of the student booklets, the items in the assessment pool are first grouped

into blocks of items. These then become the building blocks from which the student booklets are assembled. The entire item pool at each grade is divided into 14 blocks of mathematics items and 14 blocks of science items. The 28 blocks of items are distributed across 12 student booklets. To enable linking between booklets, each block appears in two, three, or four different booklets. Each student completes one booklet consisting of six blocks of mathematics and science items. Translation of the assessment was based on blocks rather than booklets. Countries translated each block once and entered the translated text into the electronic file for the appropriate test booklets.

Translators were permitted to adapt the text as necessary to make unfamiliar contextual terms culturally appropriate. Acceptable adaptations included changes in the names of seasons, people, places, animals, plants, currencies, etc. Exhibit 4.2 shows a list provided to translators detailing the types of adaptations that were acceptable.

Exhibit 4.2 Types of Acceptable Cultural Adaptations

Type of Change	Specific Change from:	Specific Change to:
Punctuation/Notation	decimal point	decimal comma
	place value comma	space
Units	centimeters	inches
	liters	quarts
	ml	mL
Proper nouns	Ottawa	Oslo
	Mary	Maria
Common nouns	robin	kiwi
	elevator	lift
Spelling	center	centre
Verbs (not related to content)	skiing	sailing
Usage	Bunsen burner	hot plate

Translators were allowed to change terms and expressions that were not familiar in their national culture, as long as the change would not affect the substance of the item. It was important, however, that translators not change any of the following when they modified the text of an item:

- the meaning of the item;
- the reading level of the text;

- the difficulty level of the item; and
- the likelihood of another possible correct answer for the item.

Although item writers and reviewers attempted to write and select items that would readily translate into the languages of the participating countries, occasionally an item proved problematic for translators. In those instances, the International Study Center was to be notified and a corresponding statement was to be included in the NRC Survey Activities Report.

4.2.5 Review of Independent Translations for Consensus

After the two translations were completed, they were compared item by item, and any differences were reconciled. In most cases, by discussing the differences in the translations of a particular item, the translators were able to agree on the version that was most appropriate for the study. A third translation expert was to be contacted if any disagreement in the translation remained.

4.2.6 Documentation of Cultural Adaptations

After a single translation had been agreed upon, the Cultural Adaptation Form was used to record all adaptations made to the achievement and questionnaire items during translation. The description of each adaptation included the international (English) term, the translated term for test items or the adapted term for questionnaire items, and an explanation of why that term was used. Translators also noted if there were any other changes in the translation. This documentation was used during translation verification, and during the achievement item analysis and review where necessary, to evaluate the quality of the translations.

4.3 Verification of Instrument Translations

Each translation went through a rigorous verification process that included verification by an international translation company, review by the International Study Center, verification of the item translations at the national centers and a check by International Quality Control Monitors.

4.3.1 International Verification of the Translations

After the final translated version of each instrument was developed, the translation was checked through an external verification process. The IEA Secretariat developed and managed the translation verification process working closely with two international translating companies with reputations for excellence, Bowne Global Solutions (formerly Berlitz), based in Luton, England, and Capstan, based in Louvain-le-Neuve, Belgium. Bowne and

Capstan staff were to document all errors and omissions and make suggestions for improvements so that National Research Coordinators could revise and improve their instruments.

Translators selected by Bowne and Capstan to serve as translation verifiers for TIMSS were required to have first-language experience in the target language, formal credentials as translators working in English, and to live and work in the target country. When the last condition could not be met, verifiers were expected to maintain close contact with the country and its culture.

4.3.1.1 Submission of Instruments for Verification

NRCs were required to send (no later than six weeks before printing) the following instruments for each grade assessed to the IEA Secretariat in preparation for external translation verification:

- one copy of the test blocks of achievement items (14 blocks of mathematics items and 14 blocks of science items) and the accompanying instructions for students;
- one set of the assembled test booklets (booklets 1 through 12); and
- one copy of the student questionnaire, teacher questionnaire(s),³ and school questionnaire.

All countries that participated in the TIMSS 2003 data collection submitted national versions of instruments for translation verification (see Exhibit 4.1).

4.3.1.2 The Translation Verification Process

The primary task of translation verifiers was to evaluate the accuracy of the translation and layout of the survey instruments. Verifiers were asked to make recommendations for improvements in the translation, when necessary, and also to alert the national centers to any deviation from the international version in the layout of the translated instruments.

Verifiers were provided with general information about the study and the design of the instruments. They also received materials describing the translation procedures used by the national centers and cultural adaptations deemed acceptable, along with detailed instructions for reviewing the instruments.⁴ The verification guidelines emphasized the importance of maintaining the meaning, difficulty level, and format of each item while allowing for cultural adaptations as necessary.

³ As noted above, at fourth grade there is one teacher questionnaire, and at eighth grade there are separate mathematics and science teacher questionnaires.

⁴ Materials provided to verifiers included *Guidelines for the Translation Verification of the TIMSS 2003 Main Survey Instruments (TIMSS, 2001)*.

Each verifier received a package consisting of:

- the international version of each survey instrument (test blocks, test booklets, and background questionnaires);
- a set of the translated national instruments to be verified, along with the Cultural Adaptation Forms;
- a copy of the instructions given to the translators in each country;
- guidelines for translation verification, including instructions for verifying the content and layout of the survey instruments and the instructions to students;
- translation verification control forms to be completed for each instrument; and
- translation verification report forms to be completed for each instrument.

For TIMSS 2003 countries that also participated in prior cycles of the study, verifiers were responsible for ensuring that the translated version of the trend items was identical to that administered in 1995 at fourth grade and 1999 at eighth grade. Accordingly, verifiers reviewing instruments for trend-study countries also received the following:

- the translated trend items used in that country in 1995 for fourth grade and/or 1999 for eighth grade; and
- a trend item verification form.

In addition to receiving detailed written instructions, verifiers had the opportunity to discuss with the IEA coordinator any problems they encountered while performing their task.

4.3.1.3 Translation Verification Reports

Two types of reports were written by the translation verifier to document the verification process. First, the translation verifier completed a translation verification control form for each instrument. This cover sheet served as a checklist indicating which materials had been verified and whether or not deviations were found in the instruments, and including the verifier's opinion about the general quality of the translation. Second, where in the judgment of the verifier the translated version of an achievement or questionnaire item deviated from the international version, the translation verifier completed a translation verification report form with entries made indicating:

- the location of the translation deviation (page and item number);
- the severity of the deviation (using a severity code as defined below);
- a description of the change; and
- a suggested alternative translation.

These records were used to document the quality of the translations and the comparability of the testing materials in each country.

The severity codes ranged from 1 (serious error) to 4 (acceptable adaptation).⁵ The severity codes were:

Code 1 – Major Change or Error: Examples include incorrect ordering of choices in a multiple-choice item; omission of a graph; complete omission of an item; incorrect translation of text such that the answer is indicated by the question; an incorrect translation that changes the meaning or difficulty of the question; incorrect ordering of the items or placement of the graphics.

Code 2 – Minor Change or Error: Examples include spelling errors that do not affect comprehension; misalignment of margins or tabs; incorrect font or font size; discrepancies in the headers or footers of the document.

Code 3 – Suggestions for Alternative: The translation may be adequate, but the verifier suggests a different wording for the item.

Code 4 – Acceptable Changes: The verifier identifies changes that are acceptable and appropriate adaptations of the item, e.g., where a reference to winter is changed from January to July for the southern hemisphere.

The layout of the documents was also reviewed during the verification process for any changes or deviations. Exhibit 4.3 details the layout issues that were considered and checked for each survey instrument.

Exhibit 4.3 Layout Issues Considered in Verification

Layout Issues	Verification Details
Instructions	Test items should not be visible when the test booklet was opened to the Instructions section.
Items	All items should be included in the same order and location as in the international version.
Response options	Response options should appear in the same order as in the international version.
Graphics	All graphics should be in the same order and modifications should be limited to necessary translation of text or labels.
Font	Font and font size should be consistent with the international version.
Word emphasis	Word emphasis should remain the same as in the international version. If the form of emphasis was not appropriate for the given language, an acceptable alternate form of emphasis should have been used (e.g., italics instead of capital letters).
Shading	Items with shading should be clear and text legible.
Page and item identification	Headers and footers that include booklet, page, and item identification should be present.
Pagination	Page breaks should correspond with the international version of the instruments.

⁵ When in doubt as to the severity of the deviation, verifiers used code 1.

If the layout of an instrument differed in any way from the international version, an entry was made in the translation verification report form indicating the location of the deviation, the severity of the deviation, and a description of the change in the layout. If necessary and appropriate, a suggestion for improving the layout was included.

For countries that participated in prior cycles of TIMSS, verifiers also completed a trend item verification form, indicating whether there was any difference in translation or format of the trend items between the 2003 version and the 1995 version for fourth grade and 1999 version for eighth grade, with a description of the nature of the change.

The completed translation verification forms were sent to NRCs and an additional copy was sent to the International Study Center at Boston College and the IEA Data Processing Center (DPC) in Hamburg, Germany. The NRCs were responsible for reviewing the reports and revising the instruments, at their own discretion, based on the translation verifiers' suggestions.

Although generally countries complied very well with the requirements for translation verification, a number of countries did not submit for verification instruments in languages that were used. Bahrain did not submit its English version of instruments for review; Egypt did not submit its English and French versions of instruments, which were borrowed from Lebanon, for review; Hong Kong did not submit its English version of achievement tests for review; Latvia did not submit its Russian version of instruments (which were borrowed from the Russian Federation) for review; the Palestinian National Authority did not submit its English version of instruments for review; Romania did not submit its Hungarian version of instruments (which were borrowed from Hungary and not adapted) for review; and Syria did not submit its background questionnaires for review. The following countries submitted test booklets but not blocks or test blocks but not booklets for review: Argentina, Armenia, Bahrain, Cyprus, Egypt, England, Lebanon, Moldova, Morocco, Palestinian National Authority, Saudi Arabia, Scotland, Syria, Tunisia, United States, and Yemen.⁶ The following countries did not submit Cultural Adaptations Forms along with their instruments for review: Bahrain, Cyprus (tests), Egypt, Indonesia (questionnaires), Japan, Jordan, Latvia (tests), Lebanon, Lithuania (tests), Morocco, Syria, Tunisia, and Yemen.

⁶ Due to time limitations, southern hemisphere countries (Australia, Botswana, Chile, Malaysia, New Zealand, Singapore, South Africa) were required to submit only the test blocks and not the test booklets to the IEA Secretariat for review.

4.3.2 International Study Center Review

For a final review, NRCs were required to submit a print-ready copy of the achievement test booklets and questionnaires to the TIMSS & PIRLS International Study Center at Boston College, after implementing the suggestions of the translation verifiers.

For all countries, achievement and questionnaire items were compared with the international version to identify any changes in text, graphics, and format, and the test booklets and questionnaires were reviewed to identify any changes in layout. The text was reviewed for format, and items were checked to ensure that they had identical translations in the stem and options across different booklets.

For trend countries, each trend item was compared to the 1995 translated version for fourth grade and the 1999 translated version for eighth grade to note if any change had been made. When the language of these items was not familiar to the reviewer, the NRC was asked about any apparent changes.

NRCs were provided with a list of any deviations identified by the International Study Center that went beyond those recorded in the translation verification reports. NRCs used these comments to correct errors prior to printing, again at their own discretion. Countries that did not allot enough time for this step of the translation and review process were not required to submit their instruments to the ISC prior to printing, so as not to jeopardize their schedule for administering the assessment. The following countries submitted their instruments to the International Study Center for final review after printing: Armenia, Bahrain, Egypt, Japan, Korea, Lebanon, Morocco, Palestinian National Authority, Slovenia, Syria, Yemen, Ontario, and Quebec. Although the Philippines submitted instruments for review prior to printing, no corrections based on IEA or ISC review were implemented prior to printing.

4.3.3 Verification of Translations at National Centers

The results of statistical item analyses from the TIMSS 2003 field test, conducted during April through June of 2002, were reviewed by each country. Since unusual item statistics could be an indication of errors in translation, each NRC was asked to check the results to identify items that might have been mistranslated.

4.3.4 International Quality Control Monitor Item Review

As part of an ambitious quality control program, International Quality Control Monitors (QCMs) were hired to document the quality of the TIMSS 2003 assessment in each country (see Chapter 7 for a description of the work of the Quality Control Monitors). An important task for the QCMs was to review the translation verification reports for each test language and verify whether the suggested changes were made in the final instruments. The QCM marked on a copy of the translation verification report form whether the change suggested in the report was implemented. This assisted the International Study Center in identifying changes made or not made to the national versions.

4.4 Summary

The rigorous procedures for translation, cultural adaptations, translation verification, and review of the instruments implemented for TIMSS 2003 provided for comparable translations of the instruments across participating countries. The verification process of internal review, external translation verification by bilingual judges, and review by the International Study Center and Quality Control Monitors proved to be a comprehensive program for verification, ensuring accuracy in the analysis and reporting of the TIMSS 2003 data.

References

TIMSS (2002a), *TIMSS 2003 Survey Operations Manual*, prepared by the TIMSS & PIRLS International Study Center, Chestnut Hill, MA: Boston College.

TIMSS (2002b), *Guidelines for the Translation Verification of the TIMSS 2003 Main Survey Instruments*, prepared by the IEA Secretariat, Amsterdam: The Netherlands.