

Chapter 5: TRANSLATION AND VERIFICATION OF THE SURVEY MATERIAL

INTRODUCTION

This chapter explains the procedures used for translation, adaptation, and verification for paper-based (Strands A and B) materials in PISA for Development.

One of the important aspects of quality assurance in PISA is to ensure that the instruments used in all participating countries to assess students' performance provide reliable and comparable information. In order to achieve this, strict procedures for the localisation (translation, adaptation, and validation) of national versions of all survey instrumentation were implemented in PISA for Development as they were in all rounds of the main PISA assessment.

Source versions of the different documents and instruments were distributed in English and French, and reference versions were distributed in Spanish. In PISA terminology, only the English and French versions have the status of a "source" version, meaning they go through a more complex process involving linguists and domain experts trained to determine an acceptable translation. A source version can be the starting point for a translation into a different version, for example, the Khmer version for Cambodia could be translated from the French or English source version. Any other centrally produced version or a previously verified version from another country would be referred to as a "reference" version. Even though in PISA-D there were no procedural differences between the production of the French/English source versions and the Spanish reference versions, it would not be acceptable to translate from Spanish into another language

The procedures for translation and verification of the instruments included:

- optimising the English source version for translation through translatability assessment (this was done only for new questionnaire items developed for PISA-D);
- developing a French source version and a Spanish reference version for items for which these did not yet exist;
- implementing a double translation design;
- preparing detailed instructions for the localisation of the instruments for the Field Trial and for review for the Main Survey;
- preparing translation/adaptation guidelines;
- training national staff in charge of the translation/adaptation of the instruments; and

- validating the translated/adapted national versions by conducting verification of the translations by independent verifiers, review by cApStAn staff and the Translation Referee or the Questionnaires team, countries' post-verification review, and “technical” and linguistic final checks.

DEVELOPMENT OF THE FRENCH SOURCE VERSION

Since the inception of the survey, it has been a requirement that the international contractor should produce an international French source version of the data collection instruments in addition to the English version. In PISA for Development there was no new item development for cognitive assessment: Most units were borrowed either from main PISA or from PISA for Schools, PIAAC (Programme for the International Assessment of Adult Competencies) or LAMP (Literacy Assessment and Monitoring Programme).

The process for producing the French source version depended on the origin of the test unit or questionnaire item. There were three possible scenarios:

- 1) Units or items coming from PISA with or without any adjustments. In this case the existing French source was used “as is” or updated to match the PISA-D English source, if necessary;
- 2) Units coming from other surveys where a verified French national version was available (but no French source was available). In this case the existing French translation was thoroughly reviewed and updated as necessary to match the PISA-D English source; and
- 3) Units coming from other surveys where no verified French national or source version was available. In this case a French source version was produced through a complex design including double translation and reconciliation.

For the context questionnaires, only the first approach was used even though some of the questionnaire items originated from surveys other than PISA such as SACMEQ (Southern and Eastern Africa Consortium for Monitoring Educational Quality) and the Canadian National Longitudinal Survey of Children and Youth (NLSCY).

The procedure used to produce the French source version of PISA-D test units, for which no French translation was available, was similar to what has been used in all previous PISA administrations. The English-to-French translation process has proven to be very effective in detecting residual errors overlooked by the item developers, and in anticipating potential problems for translation in other languages. For PISA-D, the list of aspects requiring national adaptations could be refined and further translation notes could be added as needed.

The French source version was produced through the double translation and reconciliation process, followed by a review by a French domain expert, for appropriateness of the terminology, and by a native professional French proofreader for linguistic correctness. In addition, an independent verification of the equivalence between the final English and French

versions was performed.

For the context questionnaires, the situation was similar to the cognitive items: For those questionnaire items that were adopted from PISA, the French source version could also be adopted and updated where needed. Any new questionnaire items were produced through the double translation and reconciliation design.

DEVELOPMENT OF SPANISH REFERENCE VERSION

In PISA there has only been a Spanish reference version of the cognitive items since PISA 2015, so the situation was different from French:

- 1) Units from main PISA prior to 2015, with or without adjustments: The verified Chilean version was used “as is” or updated to match the PISA-D English source, if necessary;
- 2) Units from main PISA 2015, with or without adjustments: The Spanish reference version was used “as is” or updated to match the PISA-D source, if necessary;
- 3) Units from other surveys where a verified Spanish version was available: Existing Spanish translation was reviewed and updated if necessary to match the PISA-D English source;
- 4) Units from other surveys where no verified Spanish version was available: A Spanish translation was produced through the double translation and reconciliation process.

For questionnaires there was no existing Spanish reference version in PISA. Therefore for all existing questionnaire items, the verified Chilean national version was used as the base version and updated where needed (scenario 1 above). Any new questionnaire items were produced through the double translation and reconciliation process.

DOUBLE TRANSLATION FROM TWO SOURCE LANGUAGES

Back translation has long been the most frequently used way to ensure linguistic equivalence of test instruments in international surveys. It requires translating the source version of the test (generally English language) into the national languages, then translating them back to English and comparing them with the source language to identify possible discrepancies. A second approach is a double translation procedure (i.e., two independent translations from the source language[s], and reconciliation by a third person). This offers two significant advantages in comparison with the back translation design:

- Equivalence of the source and target versions is obtained by using three different people (two translators and a reconciler) who all work on both the source and the target versions. In a back translation design, by contrast, the first translator is the only one to simultaneously use the source and target versions.
- Discrepancies are recorded directly in the target language instead of in the source

language, as would be the case in a back translation design.

Both back translation and double translation processes have a potential disadvantage in that the equivalence of the various national versions depends exclusively on their consistency with a single source version (in general, English). In particular, one would wish the highest possible semantic equivalence since the principle is to measure access that students from different countries would have to a same meaning, through written material presented in different languages. Using a single reference language is likely to give undue importance to the formal characteristics of that language. If a single source language is used, its lexical and syntactic features, stylistic conventions, and the typical patterns it uses to organise ideas within the sentence will have a greater impact on the target language versions than desirable. The recommended approach in PISA therefore builds on the strengths of the double translation approach by using double translation from two different source language versions.

Resorting to two different languages may, to a certain extent, reduce problems linked to the impact of cultural characteristics of a single source language. Admittedly, both languages used in PISA share an Indo-European origin. However, they do represent relatively different sets of cultural traditions, and are both spoken in several countries with different geographic locations, traditions, social structures, and cultures.

The use of two source languages in PISA results in other anticipated advantages such as the following:

- Many translation problems are due to idiosyncrasies: words, idioms, or syntactic structures in one language appear untranslatable into a target language. In many cases, the opportunity to consult the other source version may provide hints at solutions.
- The desirable or acceptable degree of translation freedom is very difficult to determine. A translation that is too faithful to the original version may appear awkward; if it is too free or too literary, it is very likely to jeopardise equivalence. Having two source versions in different languages, with clear guidelines on the amount of translation fidelity/freedom, provides national reconcilers with accurate benchmarks in this respect, which neither back translation nor double translation from a single language could provide.

As in main PISA, the double translation and reconciliation procedure was a requirement for all national versions of test and questionnaire instruments used in the assessment that were not adapted from the English or French source, or from the Spanish reference version. It was possible for countries to use the English source version for one of the translations into the national language and the French source version for the other. An efficient alternative method was to perform double translation and reconciliation from one of the source languages, and extensive cross checks against the second source language. Countries had to fill in and submit a translation plan documenting national translation and adaptation procedures for translation

referee review and approval.

In PISA-D, the only country that produced its instruments through the double translation and reconciliation design was Cambodia. All other countries were able to adapt their national version from one of the source or reference versions.

PISA-D TRANSLATION AND ADAPTATION GUIDELINES

PISA Translation and Adaptation Guidelines were produced to guide the national teams in the adaptation work of the instruments. The guidelines included:

- instructions on double or single translation. Double translation (and reconciliation) was required for test and questionnaire materials, but not for manuals, coding guides and other logistic material. In countries where the NPM had difficulty appointing competent translators from both French and English, double translation from English or French only was considered acceptable; in such cases it was highly recommended to use the other source version for cross-checks during the reconciliation process insofar as possible
- instructions on recruitment and training of translators
- security requirements on how to store and exchange the materials
- references to other documents, including technical guides for translating and reconciling materials
- recommendations to avoid common translation difficulties
- instructions on how to adapt the test material to the national context
- instructions on how to translate and adapt questionnaires and manuals to the national context

In addition to the generic translation and adaptation guidelines, the translators and reconcilers were given item-specific guidelines within the monitoring sheets that accompanied the materials throughout the localisation process. These guidelines provided help for specific translation and adaptation challenges. The item-specific guidelines used in the surveys where the items had been previously used were taken for PISA-D if applicable. Additional guidelines were added where needed after a thorough review of the source materials.

TRANSLATION TRAINING SESSION

NPMs received sample materials to use when recruiting national translators and training them at the national level. The NPM Meeting held in January 2016 in Rockville, Maryland, USA, included sessions on the Field Trial translation/adaptation activities in which recommended translation procedures, PISA Translation and Adaptation Guidelines, and the verification process were presented in detail separately for each component of the survey (context questionnaires and test materials).

PARTICIPATING COUNTRIES, TESTING LANGUAGES AND TRANSLATION/ADAPTATION PROCEDURES

Seven countries participated in the Strand A and B components of PISA for Development, each of them administering the test in one language. Table 5.1 below has a list of the participating countries indicating the language of the test, and translation/adaptation process used.

Table 5.1 **Participating countries, testing languages and translation/adaptation process**

Country	Testing language	Translation/Adaptation process
Cambodia	Khmer	Double translation and reconciliation from ENG source only
Ecuador	Spanish	Adaptation from ESP reference version
Guatemala	Spanish	Adaptation from ESP reference version
Honduras	Spanish	Adaptation from ESP reference version
Paraguay	Spanish	Adaptation from ESP reference version
Senegal	French	Adaptation from FRA reference version
Zambia	English	Adaptation from ENG reference version

For the coding guides, the recommended procedure was to single translate from one source version with cross-checks against the other version.

It turned out that only one country (Cambodia, for Khmer) needed to translate the material. All other participating countries adapted one of the source versions (English or French), or the Spanish reference version. Cambodia benefited, however, from the specific item guidelines linked to the development of the French source version.

INTERNATIONAL VERIFICATION OF THE NATIONAL VERSIONS

As in main PISA, one of the most important quality control procedures implemented to ensure high quality standards in the translated assessment materials for PISA for Development was to have an independent team of expert verifiers, appointed and trained by the Consortium, verify each national version against the English and/or French source versions. This quality

control step was performed for all national versions, regardless of whether they were adapted from one of the source versions, from the Spanish reference version, or double translated from English and/or French.

The main criteria used to recruit verifiers of the various national versions were that they had:

- native command of the target language;
- professional experience as translators from English and/or French into their target language;
- when possible, sufficient command of the second source language (either English or French) to be able to use it for cross-checks in the verification of the material (note that not all verifiers were proficient in French, but this was mitigated by the fact that the cApStAn reviewer and the Translation Referee had command of French);
- when possible, familiarity with the localisation of assessment materials and questionnaires;
- a good level of computer literacy; and
- when possible, experience as teachers and/or higher education degrees in psychology, sociology or education.

Prior to the verification of the Field Trial materials, training webinars were organised for the verifiers. These training sessions focused on:

- presenting verifiers with PISA objectives and structure;
- familiarising them with the material to be verified, the verification procedures, and the tools to be used;
- discussing the Translation Guidelines and the Verification Checklist;
- conducting hands-on translation exercises;
- arranging schedules and dispatch logistics; and
- security requirements.

Verification procedures have been continually improved throughout each PISA round, based on the experience and learning from previous rounds. In PISA 2015 the change from paper- to computer-based delivery mode also brought changes in the procedures; these were further refined for PISA 2018. In the following subsections we review the procedures implemented in PISA-D, based on PISA 2018 procedures, for the different components subject to verification.

VERIFICATION OF TEST UNITS (TRANSLATED AND ADAPTED VERSIONS)

For the translated versions, the verifiers were instructed to verify the target version sentence by sentence, comparing to the English or French source, while consulting the TAS (Test Adaptation Spreadsheet) to see item-specific guidelines and National Centre comments. They made changes as needed using the tracked-changes feature in Microsoft Word, documenting

their interventions in the TAS, including selection of the appropriate intervention category using a drop-down menu.

For the versions adapted from English, French, or Spanish, the verifiers were instructed to verify the changes made by the National Centre versus the reference version, and to ensure that they were:

- in compliance with the general and item-specific guidelines;
- equivalent to source version(s);
- linguistically correct; and
- correctly and consistently implemented.

It was also part of the verifiers' task to check that the errata (errors identified in source) had been corrected in the target version.

Once items from a domain were verified and reviewed, they were made available to the Translation Referee. The referee would then go through each verifier comment and label as "requires follow-up" any crucial issues that could potentially affect equivalence or item functioning. Changes labeled as "requires follow-up" were negotiated between the Consortium referee and the National Centre. The National Centre then uploaded revised Word files on the SharePoint site for final check. The final-check reviewer checked the correct implementation of any changes "requiring follow-up" and either released the files for layout check and national version construction by the Consortium or released them back to the National Centre for additional corrections.

Since the PISA 2003 Main Survey, the central element and repository of the entire translation, adaptation, and verification procedure for test units has been the TAS. Figure 5.1 shows a sample Test Adaptation Spreadsheet from the PISA for Development Field Trial. The aims of the TAS are to function as:

- an aid to translators, reconcilers, and verifiers (through the increasing use of item-specific translation/adaptation guidelines);
- a centralised record of national adaptations, verifier corrections, and suggestions;
- a way of conducting discussions between the National Centre and the Translation Referee;
- a record of the implementation status of "requires follow-up" in test units; and
- a tool permitting quantitative analysis of verification outcomes.

■ Figure 5.1 ■

Sample of a Test Adaptation Spreadsheet (TAS) from PISA for Development Field Trial

		REVIEW <i>National Centre</i>			VERIFICATION <i>e-AP/SA/An</i>		REFEREE REVIEW <i>Béatrice</i>		REVIEW <i>National Centre</i>
ENGLISH SOURCE VERSION	ITEM-SPECIFIC TRANSLATION/ ADAPTATION GUIDELINE	BORROWED VERSION (FRA source or ESP base)	NATIONAL VERSION	COUNTRY COMMENT (DESCRIPTION + JUSTIFICATION OF ADAPTATION)	VERIFIER INTER-VENTION	VERIFIER COMMENT	TRANSLATION REFEREE COMMENT	CORRECTION STATUS	COUNTRY POST-VERIF COMMENT
TABLE OF APPROXIMATE DISTANCES (in kilometers)	Do not adapt kilometers, but ok to change spelling to "kilometres"	TABLA DE DISTANCIAS APROXIMADAS (en kilómetros)	Tabla de distancias aproximadas (en kilómetros)	We usually write the titles in lowercase.	Layout/Form at issue	Changing the formatting in the stimulus is not recommendable, as upper case is used as emphasis that would then be different in this national versions vs. the others. Therefore changed to uppercase by ver.	Agree with verifier and this adaptation was not done in Math units in batch 1. Please accept the verifier correction	REQUIRES FOLLOW-UP	Ok

MAIN SURVEY VERIFICATION OF TEST UNITS

Before the Main Survey verification, the Field Trial clusters were reassembled centrally by the Consortium to reflect the item selection and order of the clusters for the Main Survey. There was a very small number of global revisions made in the source instruments between the Field Trial and Main Survey, and they were likewise centrally implemented in the Main Survey clusters before they were made available to the countries. The global revisions were documented in the Change Request Form. This form was customised per country, and those items that showed differential item functioning based on item response theory fit statistics were flagged either as “easier than expected” or “harder than expected”. Countries were then asked to:

- check if there were any linguistic or translation issues in their “flagged” items;
- check that the global revisions (wording as well as layout/structure of files) had been correctly reflected in their national version;
- check if there were any other “known errors”, that is, errors spotted after the Field Trial instruments were finalised; and
- document any necessary change requests in the change request form.

The verification of the Main Survey instruments consisted of verifying the changes that countries requested to their Field Trial instruments, either based on poor performance or differential item functioning in the Field Trial, or on the detection of residual errors. Verifiers centrally implemented those changes that were approved by the Translation Referee. Countries did not have editing access to their units or clusters at this stage. Figure 5.2 below shows an example of a change request inserted by the National Centre in the Change Request Form.

■ Figure 5.2 ■

Sample of a documented change request to a flagged item in a Field Trial-to-Main Survey change request form

National Centre					Core A (ETS)	cApStAn	Translation Referee
ITEM ERROR and JUSTIFICATION FOR CHANGE	ENG/FRA source wording	Original FT wording	English back-translation of original FT wording	Corrected MS wording	TEST DEVELOPER COMMENT	VERIFIER COMMENT	TRANSLATION REFEREE COMMENT After verification
In the IA and IRT report we have, this item is not flagged. Could you please tell us which seems to be the trouble with this item? Also, after a new check we came to realize that 'the drawing' may be interfering with the comprehension, since both the plan and the room may seem as drawings for the respondents. Will it be possible to delete this part? Do you think it's prudent during MS? The new suggested line should read: ¿Qué plano representa de mejor manera la habitación?	Which plan best represents the drawing of the room?	¿Qué plano representa de mejor manera el dibujo de la habitación?		¿Qué plano representa de mejor manera la habitación?	In fact, this item should NOT have been flagged. I would recommend no change. I went ahead and changed the Flagged Item status to NO. (AK)	The item was not flagged, corresponds to source and is linguistically correct, so ver recommends against changing, as per TD comment in col. O.	Please see comment in O and P. As this item worked well in the FT, changing the wording could have adverse effects. However I quite agree with your explanation in I. For the reasons you give in column I, French source reads: Which plan best represents the room that is shown on the drawing. If you still think that your wording should be changed, then I would recommend to align your wording on the French source version. Please advise in column T

FIELD TRIAL VERIFICATION OF CODING GUIDES

As in main PISA, the PISA-D coding guides were verified separately from the cognitive items, and at a later time. The overall procedure was the same as for paper-based test units: Verifier corrections were made in tracked changes in the Word files and documented in the monitoring sheets in Excel format.

MAIN SURVEY VERIFICATION OF CODING GUIDES

The process for verifying the coding guides for the Main Survey was the same as for the cognitive items. Countries documented their change requests in a change request form. Unlike in the Field Trial, during the Main Survey verification took place at the same time with the cognitive items, and within the same form.

FIELD TRIAL VERIFICATION OF CONTEXT QUESTIONNAIRES

Context Questionnaires were submitted for verification together with an agreed Questionnaire Adaptation Spreadsheet (QAS). The main purpose of the QAS was to document all content-related or “structural” deviations from the international source or reference versions. Such national adaptations were subject to clearance by the questionnaire team at The Learning Bar before the material was submitted for verification. Subsequently, the QAS served the same objectives and followed the same logic as the TAS for test units (see above). Figure 5.3 shows a sample QAS from the PISA for Development Field Trial.

■ Figure 5.3 ■

Sample of a Questionnaire Adaptation Spreadsheet (QAS) from PISA for Development Field Trial

SOURCE VERSION AND TRANSLATION AND		ADAPTATION GUIDELINES STEP 1 - NEGOTIATION OF ADAPTATIONS				STEP 2 - VERIFICATION		
<i>Pre-populated by</i>		<i>National Centre to complete</i>			<i>Questionnaire Team to complete</i>		<i>Verifier to complete</i>	
International English Version	ITEM-SPECIFIC TRANSLATION ADAPTATION GUIDELINE	Proposed target version ONLY if adapted	English translation of the proposed target version ONLY if adapted	Justification for proposed adaptations; National Centre comments	Questionnaire team comments	Agreement Status	Verifier intervention category	
		Verifier comments						
Which of the following do you <u>expect</u> to complete?	Keep underlining or equivalent emphasis	¿Cuál de los siguientes niveles de educación esperas completar?			This is a translation issue to be verified by cApStAn.	<i>Agreed</i>	Register/ wording issue	"esperas" can also be interpreted as "hope", the expression is a hybrid between hoping and expecting. As this question wants to catch the student's expectation, we believe "anticipar" [¿Cuál de los siguientes niveles de educación anticipas completar?] is more to the point, since the expression involves a (probability) calculation and is not just the product of dreaming or wishing. "Tienes previsto" is also a possibility, a bit more elegant, but the meaning is the same. Also underlining was missing. Changed to "anticipar" by ver.

For the translated version (Khmer), the verifier’s brief was defined as checking whether target questionnaires were linguistically correct and faithful to either the source version (when no adaptation was made) or the approved English translation of the national version (when an adaptation was made). With a view to this, the verifier was instructed to:

- check whether the back translation of the agreed adaptation was faithful;
- check whether the agreed adaptation was correctly reflected in the questionnaire;
- check the questionnaires for undocumented adaptations (deviations from the source not listed in the QAS) and report them; and
- check linguistic correctness (grammar, spelling, etc.) of the entire target version.

Verifier interventions were entered in the questionnaires using the tracked changes mode, while verifier comments were entered in the verifier column of the QAS.

As for test units, any further significant changes were labeled as “requires follow-up” by the Questionnaire team, and after negotiation with the country teams, their correct implementation was checked by verifiers during final check.

For the versions adapted from English, French, or Spanish, the verifiers’ brief was to

- check that the changes made by the National Centre versus the reference version were linguistically correct and consistently and correctly implemented; and
- check that the agreed adaptations were correctly reflected in the questionnaire.

As for the cognitive items, it was also part of the verifiers’ task to make sure that the errata were correctly addressed in the target questionnaires.

MAIN SURVEY VERIFICATION OF QUESTIONNAIRES

In the process for preparing the questionnaire instruments for the Main Survey, the countries were asked to:

- reorganise their final Field Trial questionnaires to match the Main Survey source version;
- implement the global Field Trial-to Main Survey revisions in their national version; and
- request national Field Trial-to-Main Survey changes, if needed.

For the questionnaires there was no separate change request form. Instead the Field Trial QAS forms were reformatted prior to making them available to the countries: the order of items was changed to match the Main Survey source; items that were dropped after the Field Trial were removed; and Field Trial-to-Main Survey revisions were implemented in the “Main Survey source” column, clearly marked and explained. The countries used this reformatted Main Survey QAS to request changes.

The verifiers’ brief was to check that the global revisions were correctly implemented and that the change requests made by the country were linguistically correct, in compliance with the general and item-specific guidelines, equivalent to source, and correctly and consistently implemented. Any additional change requests beyond the global revisions were subject to approval of the Questionnaire team at The Learning Bar.

SUMMARY OF ITEMS DELETED AT THE NATIONAL LEVEL, DUE TO TRANSLATION, PRINTING, OR LAYOUT ERRORS

In the Main Survey, no items were dropped at the national level related to translation, printing, or layout errors.