

PISA 2022 Technical Report



7

Translation and Verification of the Survey Material

Introduction

This chapter describes the translation and adaptation procedures, the linguistic quality control (verification) procedures for both paper-based (PB) and computer-based (CB) materials in PISA 2022, as well as the upstream linguistic quality assurance procedures used to produce the source versions of the PISA instruments.

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. To achieve this, strict procedures for the localisation (adaptation, translation, and validation) of national versions of all survey instrumentation were implemented in PISA 2022, as in all previous PISA rounds.

These procedures included upstream and downstream linguistic quality assurance processes, further explained below.

Upstream linguistic quality assurance Processes include the following aspects:

- Optimisation of the English source version for translation through translatability assessment.
- Development of two source versions of the instruments, in English and French (except for the Financial Literacy and for the operational manuals, provided only in English).
- Implementation of a double translation design with a final reconciliation.
- Preparation of detailed instructions for the localisation of the instruments for the Field Trial and for their review for the Main Survey.
- Preparation of translation/adaptation guidelines.
- Production of item-by-item translation and adaptation notes.
- Training of national staff in charge of the translation/adaptation of the instruments.
- Centralised trend material transfer.

Downstream Linguistic Quality Control Processes include the following aspects:

- Validation of the translated/adapted national versions: verification 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.
- Centralised management of the changes and updates in the trend materials.

PISA Countries/economies, Languages, Scope and Verifier training

The countries or economies participating in PISA 2022, referred to in this report as PISA Countries/economies, were responsible for the translation and adaptation of their instruments. Table 7.1 lists the verified language versions with the following additional information:

- ISO (three-letters) Code 3366
- The last cycle in which they participated in PISA
- The mode of administration (PB for Paper Based or CB for Computer Based assessment)
- The change of mode compared to the last cycle (paper-based assessment (PBA) → CBA)
- Whether the version was adapted from the English or French source, from the common base version in Spanish or Chinese, or from borrowed version from another country/economy
- The international options that underwent the verification process: CT (Creative Thinking), FL (Financial Literacy), ICQ (Information and Computer Literacy Questionnaire), TCQ (Teacher Questionnaire), WBQ (Well-being Questionnaire), UH (Une-heure test and questionnaires), PAQ (Parent Questionnaires).

While most of the PISA 2022 countries/economies has also administered the assessment in PISA 2018, five countries/economies with six versions were new to PISA 2022: El Salvador, India with two languages, Hindi and English, Jamaica, Mongolia and Uzbekistan. In total, 113 language versions in 54 languages for 86 PISA Countries/economies were verified in PISA 2022. The table does not include minority language versions that represented less than 10% of the target population and were not centrally verified.

Materials subject to verification

The following materials were subject to international verification before the Field Trial:

Cognitive units

The PISA 2022 cognitive assessment consisted of the units from the three core domains, compulsory for all the PISA Countries/economies, and the international options. These include the following units:

New Mathematics units

Mathematics was the main domain in PISA 2022: 61 of newly-developed Mathematics units were translated and verified in three batches. In past cycles, the PISA Countries/economies administered one of the two “easy” or “hard” clusters. Both clusters were administered by all the PISA Countries/economies in PISA 2022, referred to as cluster 6A and 6B. Either 6A or 6B cluster was verified as new for all the countries/economies. New Mathematics units were computer-delivered and were translated and verified in XLIFF format (tagged XML Localisation Interchange File Format) in the open-source CAT (computer-assisted translation) tool OmegaT. The units were released in 4 batches for translation and adaptation, with 6A or 6B cluster in a separate batch. See Table 7.2.

From this pool, 16 new units and 10 new items were dropped for the Main Survey. See Table 7.3.

Financial Literacy units

Three new Financial Literacy units were added to the cognitive assessment pool for PISA 2022, translated, verified, and administered by the countries/economies that have chosen this international option in Field Trial and Main Survey.

Creative Thinking units

Creative Thinking was the new domain introduced in PISA 2022 as an international option, with 21 units. Unit T54, Infographics was not administered in the Main Survey. See Table 7.4.

Mathematics units (trend)

45 trend units were administered in the Field Trial, from which 2 units and 3 items were dropped for the Main Survey.

Financial Literacy units (trend)

Three new Financial Literacy units were added to the test pool for PISA 2022, translated, verified, and administered by 21 countries/economies that have chosen this international option in Field Trial and Main Survey in 28 national versions.

Reading units (trend)

Forty-nine Reading Literacy trend units were administered in the PISA 2022 Field Trial and Main Survey. For the countries/economies new to PISA, the trend Reading units were translated and verified as 'new' materials following same workflow and procedure as for new Mathematics units.

Science units (trend)

Twenty-four units from the trend Science instruments were administered in PISA 2022 Field Trial and Main Survey. Like Reading units, countries/economies new to PISA followed the workflow and procedures same as for the new Mathematics units.

Orientation, Help, Interface and Test flow files (XYZ files)

There was one new 'orientation' and one new Help file verified for all CBA countries; orientation file for FL was verified for countries/economies taking those options. The Creative Thinking orientation file was translated and verified with the CT units.

Orientation, Help, Interface and Test flow files (XYZ files) (trend)

There were nine files with other widgets, or "XYZ files", included interfaces for the calculator and Math editor, generic navigation elements, a help file, the interface for the test environment orientation files for the questionnaires, Reading, and test flow. The new PISA Countries/economies that administered the units on computer, translated these files in OmegaT, and they were all verified.

Paper-based clusters

For countries/economies administering PISA 2022 as a paper-based assessment (PBA countries/economies), the cognitive test consisted of trend units only, as no new PB items were developed for PISA 2022. For countries/economies that were new to PISA, all 44 Math, 32 Science, 22 Reading units and 4 Reading components were treated as 'new' materials and underwent the translation and/or adaptation process.

Contextual Questionnaires

There were two required contextual questionnaires, administered by all participating countries, and five optional questionnaires:

Required Questionnaires

- School Questionnaire (SCQ) with 83 questions administered on the Questionnaire Adaptation Tool (QAT) for CBA countries; for PBA countries 69 questions were translated and verified in Main Survey Word format, and administered on paper;
- Student Questionnaire (STQ) for PBA countries was administered in paper-based format (MS Word) in two booklets, each of them consisting of 15 Core questions, identical between the two booklets, as well as 30 additional questions in Booklet 1 and 42 additional questions in Booklet 2. The CBA countries administered the Student Questionnaire with 168 questions in the QAT.

The Global Crisis Module (GCM) were questions added in SCQ and STQ following the outbreak of the COVID-19 pandemic. Counts of GCM questions are included in the counts above.

Optional Questionnaires

- Parent Questionnaire (PAQ) with 45 questions available in paper-based format for both PBA and CBA countries. The Parent Questionnaire was verified in 13 languages (corresponding to 20 national versions) in 17 countries, all of these CBA countries. No PBA country opted for the Parent Questionnaire.
- Information and Communication Technology Questionnaire (ICTQ) with 14 questions administered in the QAT (70 versions verified for 57 CBA countries);
- Teacher Questionnaire (TQ) with 77 questions included in the QAT (24 versions verified for 20 CBA countries). Some questions were addressed specifically to mathematics teachers.
- Financial Literacy Questionnaire (FLQ) with 14 questions included in the QAT (31 versions for 23 countries) by countries that also opted for the Financial Literacy cognitive assessment.
- WBQ with 25 questions included in the QAT (21 versions verified for 16 CBA countries).

Verifier qualifications, training and instructional materials

As in previous PISA cycles, one of the most important quality control procedures implemented to ensure high-quality standards in the translated assessment materials was to have an independent team of expert verifiers, appointed and trained by the international contractors, verifying each national version against the English and/or French source versions.

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,
- if 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 are proficient in French, but this is mitigated by the fact that the cApStAn reviewer and the translation referee have command of French,
- if possible, familiarity with the main domain assessed,
- a good level of computer literacy and experience with computer-aided translation tools (CAT tools),
- if possible, experience as teachers and/or higher education degrees in psychology, sociology, or education.

All verifiers were invited to attend one of the two seminars, based on the verification schedule of their country. In total 32 verifiers of early-testing countries and 10 members of the cApStAn team attended the first training seminar in June 2019, and 20 verifiers and 10 cApStAn team members the second training

seminar in September 2019. A 2-day verifier training seminar was organised by cApStAn in Brussels on 31st May and 1st June 2019. In total 55 verifiers and 10 members of the cApStAn team attended the seminar. Those verifiers who were not able to come to the seminar were trained through remote Webinars in July and/or August of 2019.

The main aim of the training was to provide verifiers with background information on PISA 2022 in general, and on the verification task in particular. Verifiers were divided into four different groups based on two criteria (experienced/new and full verification/focused verification process) to attend parallel sessions:

- **Experienced verifiers** – verifiers who had participated in previous PISA cycles and had already acquired experience in verifying PISA materials.
- **New verifiers** – verifiers who had been recruited for this cycle of PISA.
- **Verifiers of adapted versions** – verifiers verifying a version adapted from the French or English source version, from the Spanish or Chinese common base version, or from a verified version produced by another National Centre.

Each group participated in three sessions:

- **Cognitive Materials** – Topics for this session included: nature and new features of the new Mathematical literacy units, challenges of mathematics units compared to other domains; structure of the TAS (Test Adaptation Spreadsheet), as well as the overall verification workflow using the portal previews. The session included hands-on exercises where verifiers edited mock XLIFF files using OmegaT, previewed the resulting file on the PISA portal and documented their findings in a TAS, under the supervision of the cApStAn trainers. The session for new verifiers' group included a generic part explaining the essence of the verification task and more background information on the PISA survey, while this was omitted in the presentation for experienced PISA verifiers. Similarly, the session for verifiers of adapted versions focused on what is relevant for this procedure, drawing examples from adapted versions in previous cycles.
- **Questionnaires** – In this session, the differences in procedure and focus of questionnaire verification vs. verification of cognitive materials was explained. There were also hands-on exercises, where verifiers were asked to work in the Questionnaire Adaptation Spreadsheet (QAS) and on OmegaT, and to verify mock translations.
- **Documentation and tools** – This session concentrated on the principles of documenting verification outcomes using the verifier intervention categories (See Annex 7.A) in a way that is informative, concise, and useful to all parties involved. Examples from previous cycles were discussed among the group to illustrate best practices in comment writing.

Tailoring the sessions to smaller groups proved to be effective in the PISA 2015 and PISA 2018, so the same approach guided the organisation of the trainings for PISA 2022.

Day 1 of the seminar was devoted to OmegaT. During the morning plenary session, the CAT tool and its features were introduced. The group was then split in parallel sessions to give the verifiers the opportunity to perform some practical exercises in smaller groups. A specific meeting for verifiers of right-to-left languages was also organised. At the end of the day, the groups were reunited for a general question-and-answer session.

Day 2 included the following sessions:

- **General PISA session** – Overview of the PISA 2022 Field Trial.
- **Cognitive Materials** – Topics for this session included: a generic part explaining the essence of the verification task and more background information on the PISA cognitive materials, the overall verification workflow, the nature and challenges of New Maths and Creative Thinking units. For the translated versions, the verifiers were divided in two smaller groups. The session for verifiers of

adapted versions focused on what is relevant for the versions adapted from one of the source versions, from a common base version or from a translation borrowed from another country.

- **Documentation and tools** – This session concentrated on the principles of documenting verification outcomes using the verifier intervention categories in a way that is informative, concise and useful to all parties involved. The novelty of the standardised comments was also illustrated. Some practical exercises were organised.
- **Questionnaires** – In this session, the differences in procedure and focus of questionnaire verification vs. verification of cognitive materials was explained. The questionnaire workflow was presented, and there were also hands-on exercises, where verifiers were asked to work in the QAS and on OmegaT, and to verify mock translations.
- **Coding guides** – In this session, the focus of verification of the coding guides was explained and the Countries/economies were explained how to take advantage of the translation memories¹ that are coming from the cognitive units. A few sample responses were shown as example.

Splitting certain sessions in smaller groups and organising hands-on exercises proved to be effective in past cycles, so the same principle was followed for PISA 2022.

Testing languages and translation/adaptation procedures

National project managers had to identify the testing languages according to the PISA technical standards and following the instructions given in the School Sampling Preparation Manual and to record them in the sampling form Sampling Task 0 (ST0) for agreement by the PISA Contractors.

In addition, based on the approved ST0, and prior to the Field Trial, national project managers had to complete a translation plan describing the procedures used to develop their national versions and the different processes used for translator/reconciler recruitment and training. Information about a possible national expert committee was also sought. This translation plan was reviewed by the translation referee for discussion/approval.

Table 7.5 summarises the Field Trial translation procedures for tests and questionnaires, as described in the confirmed translation plans. The figures in the table do not include minority language versions that represented less than 10% of the target population and were not centrally verified.²

The total number of the versions in Table 7.5 would not represent the total number of verified versions because some Countries/economies had different procedures for different domains or questionnaires, e.g., Romania double translated the cognitive units from English with cross-checks against the French source version, but for the Reading Literacy trend units that were double translated from English and French, Colombia adapted the common reference version but double translated the Parent Questionnaire from English source.

Note that for the Catalan, Galician and Basque versions, the cross-checks were made against the verified Spanish version of Spain rather than against the other source version.

Countries sharing a testing language were strongly encouraged to develop a common version in which national adaptations would be inserted or, in the case of minority languages, to borrow an existing verified version. In previous survey administrations we found that high-quality translations and high level of equivalence in the functioning of items were achieved in countries that shared a common language of instruction and could develop their national versions by introducing a limited number of national adaptations in a common version. Additionally, a common version for different countries sharing the same testing language implies that all students instructed in a given language receive booklets that are as similar as possible, which potentially reduces cross-country differences due to translation differences.

Co-operation between countries sharing the same language was therefore fostered and facilitated. To this effect, workable models were designed so that verified versions from one country could be adapted by another country.

Different scenarios of sharing were applied in the following cases:

- As in previous cycles, the model followed by German-speaking countries was highly efficient: the German version of each of the components of the assessment material was double translated and reconciled by one of the German-speaking countries, then verified, and adapted by the other countries who administered that component. The adapted versions were then verified.
- A Spanish common reference version of the new test materials was produced by an independent contractor and shared by the Spanish-speaking countries.
- A Chinese version of the new test materials was produced by an independent contractor and shared by the Chinese-speaking Countries/economies.
- A Russian common reference version was fully verified and then adapted by Azerbaijan (Baku), Estonia, Kazakhstan, Latvia, and Moldova.
- Finally, Bosnia and Herzegovina, Montenegro and Serbia shared the translation effort translating each one part of the assessment and then adapted the verified versions to their local contexts.

Development of source versions

Translatability assessment

Translatability assessment is an effort to combine linguists' expertise with that of item developers to bridge the gap between a draft item written in the source language, and an actual source version of that item, suitable for translation/adaptation.

While item writers are increasingly aware of localisation issues, they are rarely in a position to identify some of the challenges translators will be confronted with. In line with the trend to do more upstream work, i.e., work before the start of the actual translation process, a methodology was developed to identify and document potential translation and adaptation difficulties in draft PISA 2022 items before the source versions were finalised. This process, referred to as the translatability assessment, was first implemented in PISA 2015.

Translatability assessment consists of submitting draft versions of new items to a pool of experienced linguists covering a broad range of language groups. The linguists were selected among the international verifiers and were trained to use a set of 13 translatability assessment categories to report on potential translation, adaptation and cultural issues that could affect translatability.

For both new Questionnaire items and New Maths and Creative Thinking items there were always at least three linguists from different language groups evaluating each item. The approach was for each linguist to first mentally translate each item allocated to them. When the item appeared straightforward to translate, it was classified as "straightforward." When the linguist found the item somewhat difficult to translate/adapt or identified a potential cultural issue, they went through the exercise of (i) producing a written translation of that item; (ii) selecting the relevant translatability category (such as "Unnecessarily complex" or "Potential cultural issue") – see Annex 7.B; (iii) describing the issue; and (iv) proposing an alternative wording or a translation/adaptation note to circumvent the problem. It should be noted that the translations produced in category (i) were not intended for further use; they were used to help the linguists identify and describe the translation and adaptation challenges that translators might face if no pre-emptive action were taken.

The feedback from the different linguists was then collated by a senior linguist at cApStAn and reviewed by the translation referee. The senior linguist reformulated the comments so that similar issues were processed in a consistent way; selected or rewrote proposals for alternative wording that addressed all the issues identified and drafted translation/adaptation notes when applicable. When several linguists working in different languages pointed out similar issues in a given item, special attention was given to the wording of that particular item. The senior linguist produced a Translatability Report, which was then sent to the item developers for review. The item developers then used the report to eliminate ambiguities, e.g., Anglo-Saxon idiosyncrasies that may be difficult to render in certain languages, double-barrelled questions, cultural issues or unnecessary complexity. Overall, the aim was to fine-tune the initial version of the items so that it became a more translatable source version.

Production of the second source version in French

Since the inception of PISA, it has been a requirement that the international contractor should produce an international French source version of the data collection instruments. Experience has shown that some issues do not become apparent until there is an attempt to translate the instruments into a second language. As in previous PISA survey administrations, the English-to-French translation process proved to be very effective in detecting issues not perceived by the item writers, and in anticipating potential problems for translation into other languages. A number of ambiguities or pitfall expressions could be spotted and therefore avoided in the source versions by slightly modifying both the English and French source versions. As a result, the list of aspects requiring national adaptations could be refined; and further translation notes could be added as needed.

The new PISA 2022 items were first drafted in English, and then a parallel source version of the items was produced in French. The parallel source version was produced for the new Mathematical literacy items (stimuli, items, and scoring rubrics for open-ended items), the newly-developed items for the School Questionnaire, Student Questionnaire, Information and Communication Technology (ICT) familiarity Questionnaire, as well as the assessment materials for Creative Thinking (stimuli, items, and coding guides). No French source version was produced for the new Financial Literacy items.

The workflow for producing the French source was the same for newly-developed PISA 2022 Mathematics units, Creative Thinking units and Questionnaire materials. Once feedback from the translatability report and from country reviews was integrated into the revised units in XLIFF format, the translation monitoring forms in Excel format (Test Translation Spreadsheets, TTS) were prepared for the translation process into French.

There was one TTS for each batch of units and questionnaires. The form was designed to include the whole history of the process and to accommodate (i) comments from translators 1 and 2; (ii) comments from the reconciler (about FRA or about ENG source); (iii) feedback from the domain expert; (iv) consolidated feedback from the lead reconciler (about FRA or about ENG source); (v) first reactions from the test developers, (vi) issues reported during the equivalence and linguistic purity check (ELPC), (vii) second round of feedback from the item developers and (viii) proofreading at the end of the process and potential comments about residual mistakes.

In the TTS, some provisional item-per-item translation and adaptation guidelines from the TA were already included for reference and all players were invited to review these and complement with new guidelines as difficulties were identified. The final item-per-item guidelines were then used to populate the Field Trial Verification form.

The translation of the cognitive units for Mathematics and Creative Thinking was done using XLIFFs so consistency could be maximized from the very beginning of the process. The Questionnaires were received in Main Survey Word format. In PISA 2022, OmegaT was also used for the production of the questionnaires

to guarantee the same level of consistency as for the cognitive units. All materials went through a dedicated workflow on the PISA portal.

The workflow was streamlined so that the item-per-item translation and adaptation notes were formulated while the English source version was being finalised. This allowed monitoring the relevance and effectiveness of these notes early on and making necessary adjustments as the parallel source version was produced. The source version optimisation also included work with the Core A Contractor to apply segmentation rules, and to prepare style guides and rule sets for automated consistency checks.

A team of six translators, three reconcilers, two domain experts (one for the Mathematics units and one for the Creative Thinking units and Questionnaires), four equivalence and linguistic purity check reviewers and one proofreader was set up to produce the PISA 2022 French source. Most members of the team had already participated in producing the French source version of PISA 2018 instruments.

Before the start of the translation, a training workshop with all translators and reconcilers of the parallel source was held in Brussels in December 2018. All translators, reconcilers and domain experts attended the face-to-face training workshop. The training programme included a session on the translation of mathematical language and a hands-on training session to hone the translators' and reconcilers' skills in using specific computer-aided translation tools to their full potential. Sample materials from this cycle and interesting examples from the translatability assessment were used to refresh their memories, and hands-on exercises were organised to introduce the PISA portal and the tools used by cApStAn for this cycle of PISA, including OmegaT, the computer-aided translation tool. There was also an OmegaT helpdesk available throughout the translation process.

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 using the same procedures and verification checklists as for the verification of all other national versions.

The team of translators consisted of one translator who focused primarily on accuracy and systematically conveyed each piece of information in the target version, as well as one translator who concentrated primarily on fluency. As shown in Figure 7.1, the workflow began with producing the two independent translations, T1 and T2. The work was split between Questionnaires translation and cognitive item translation for Mathematics and Creative Thinking. Both translators received the same materials at the same time and delivered their translations to the reconciler on the same date.

Figure 7.1. Translation workflow for the production of a French source version of newly-developed PISA 2022 Mathematics units



Both for the new Mathematics units and the questionnaires, translation memories were created from the PISA 2018 and PISA 2015 French source of the questionnaires and added as reference. Translations of the trend questions were thus pre-populated in OmegaT, and all players were instructed to align the translations of the new questions to the trend ones. A glossary of compulsory adaptations, so called "forced

adaptations" from the previous cycles was also prepared and included in the OmegaT projects. Special attention was given to consistency across questionnaires focusing on scales, recurring instructions and forced adaptations. The translation memories from the previous cycles were useful for obtaining better consistency, especially for the questionnaires and the recurring instructions in the new Mathematics units.

The main task of the reconciler was to merge the two independent translations in such a way that the resulting national version is as equivalent as possible to the initial source version while the wording is as fluent as possible. Correspondingly, it was the lead reconciler's responsibility to finalise their single translation for the coding instructions. In particular, the reconciler ensured consistency between the French version of coding instructions and the French reconciled version of the stimuli and items, and between the English and French source versions. The lead reconciler collated the documentation on all cases where the double translation process (and single translation process for the coding guides) revealed possible flaws in the initial source version and established the communication with the item developers.

The reconciler received the OmegaT packages containing the source XLIFF files, the translation memories from the two translators (T1 and T2) and the Excel monitoring sheet for that batch. The advantage of using XLIFFs already at this stage (instead of Excel files or storyboards) was that it was possible to preview both the English and French version of the unit on the portal, so each translation could be reviewed in its real context. Another important advantage of XLIFFs is that translation of recurring elements could easily be harmonised using the translation memory utility in OmegaT. During this process, the reconciler could enter comments in the Excel monitoring sheet for the attention of the domain experts and the lead reconciler. These comments could relate to the translation and adaptation guidelines, to the English version (linguistic or contents) or to the French version. There were therefore different columns devoted to these comments. The column "Reconciler comment about ENG source" contained reconciler comments about linguistic or content issues as well as some recommendations or suggestions about the ENG wording. These suggestions were mainly aimed to improve consistency or to facilitate the translation into the different PISA languages. Suggestions for item-by-item translation and adaptation notes could also be included in this column. In the column "Reconciler comment about FRA source", the reconciler could explain some of the choices made and document issues for which the domain expert's advice was requested.

Two domain experts from France reviewed the reconciled translations of the new assessment items from the Mathematical literacy and Creative Thinking domains as well as of the new questionnaire items. The domain experts' task was to check whether the terminology was deemed appropriate for 15-year-old students; to ensure that the prompts and instructions were clear and relevant, and to evaluate whether, from their expert's perspective, the cognitive items seemed to measure the same knowledge and skills across the two languages. For the questionnaire items, the domain expert was asked to evaluate that the instruments would collect the same information in each language. The domain experts' feedback was then processed by the lead reconciler, who either implemented a change directly, or to added it to a compilation of issues that required input from the item developers at Core A and Core B3.

The feedback from the reconciler and the domain experts about the English version was then consolidated by the lead reconciler and shared with the item developers, who reacted to both the reconciler's and the domain experts' comments and provided suggestions for edits or in some cases a completely new version of the source wording in English. If a proposed change was relevant for the English master version, the updated English version was entered in the Excel monitoring sheet and the French version was then updated as needed during or after the equivalence and linguistic purity checks.

The interaction between the lead reconciler and the test and questionnaire item developers contributed additionally to the maintenance of semantic, linguistic and insofar as possible, psychometric equivalence between the two parallel source versions. The discussion between the different players was performed by documenting the issues in the TTS. Special attention was given to evaluating the impact of each edit on other parts of the materials and ensuring that the Core A and B3 item developers echo all necessary modifications in the English source.

Once the feedback from the lead reconciler and the item developers was reflected in the French source, it was submitted for a linguistic purity check and semantic equivalence check. These two checks were performed in tandem by (i) a senior staff member of cApStAn who is bilingual English/French and has expertise in the international verification of the PISA materials, who focused primarily on the finer residual equivalence issues; and ii) a native French linguist, who focused primarily on the finer points of strictly correct French language usage. The feedback from this step consisted of comments, suggestions for rewording (sometimes of the English text instead of or in addition to the French text), and proposals for translation/adaptation guidelines.

A senior cApStAn consultant processed the results of the feedback of these two steps simultaneously and shared the reports with the item developers when the reported issues had a potential impact on the English master version. This led to the second round of updates in the English source. Whenever a change in the French version was required, the final version was inserted in a specific column of the monitoring sheet, and this was then centrally transferred into the French XLIFF file by the proofreader.

Once the item developers' feedback had been implemented, a proofreader reviewed the final proofs in XLIFF format. The proofreader saw the materials for the first time in this step. This allowed them to review the final version of the French source version with a 'fresh eye', and correct residual typos, as well as grammar and syntax errors. The proofreader used the 'preview' utility on the PISA portal to proofread the materials. This allowed them to view the items exactly as the respondents would see them. When an issue was spotted, the necessary changes were made in the corresponding XLIFF; then the proofreader would refresh the preview window in order to check that the modifications were correctly implemented. The edits were limited to corrections of outright errors overlooked in the earlier steps or accidentally introduced when processing the feedback from the equivalence and linguistic purity check. The proofreader also left comments in the TTS about any residual issues identified at this step (for instance, incorrect final layout, source updates not implemented etc.) for the item developers' attention.

The coding guides for open-ended items were single translated by one of the translators from the team who produced the coding guide for the particular domain, which was first reviewed by the reconciler and the domain expert and then consolidated by the lead reconciler. Finally, the coding guides went through the equivalence and linguistic purity check process and final proofreading.

Both the translatability assessment and the development of the French source version contributed to providing national project managers (NPMs) with source material that was easier to translate and contained fewer potential translation problems than would have been the case had only one source been developed without a translatability assessment.

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 design (i.e., two independent translations from the source language(s), and reconciliation by a third person).

This second approach 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. On the other hand, in a back translation design 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 designs 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). 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 (Grisay, 2003). The recommended approach in PISA therefore builds on the strengths of the double translation approach by using double translation from two different source languages.

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 second 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 literal 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 previous PISA cycles, the double translation and reconciliation procedure were a requirement for all national versions of test and questionnaire instruments used in the assessment. 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. For the optional Financial Literacy domain, the units were double translated from English only, as there was no French source version of these units.

Training and instructional materials for national translation teams

National project managers received sample materials to use when recruiting national translators and training them at the national level. The NPM meeting held in March 2019 in Vienna 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 the questionnaires, new cognitive units, trend units and coding guides, separately for the CB and paper-based administration, and separately for the new PISA Countries/economies.

PISA 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 double translation, it was recommended that one independent translator use the English source version while the second use the French version. In countries where the National Project Manager (NPM) has difficulty appointing competent translators from 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.
- Security requirements.
- References to other documents, including technical guides for translating and reconciling CBA materials.
- Recommendations to avoid common translation traps.
- 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 were produced based on a thorough review first of the English source, then of the comments arising from the translatability assessment and then of those arising from the production of the French source version.

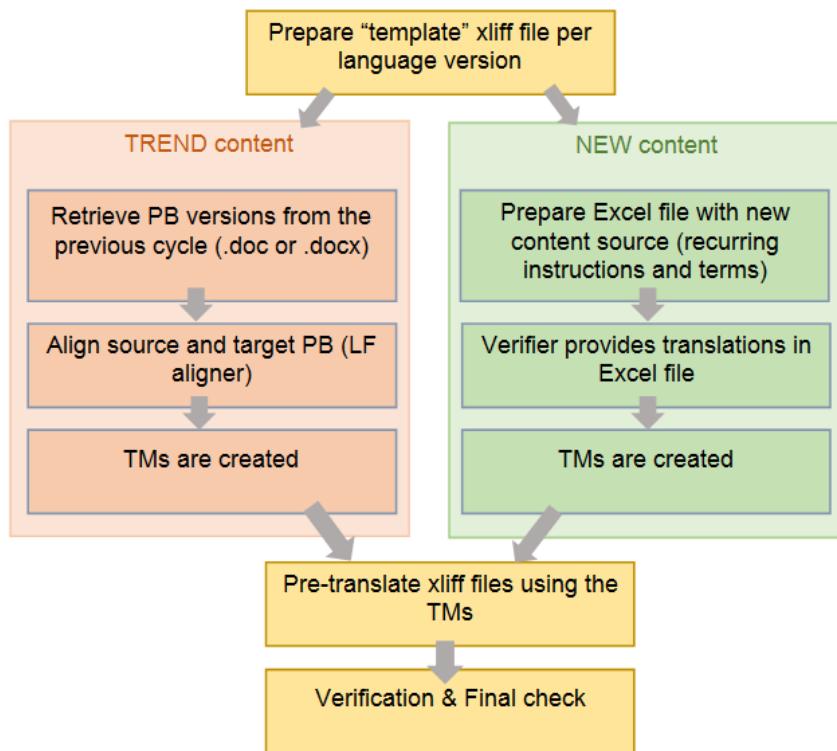
Centralised trend material transfer

Cognitive units were administered in paper-based format (MS Word) until and including PISA 2012. In PISA 2015, most participating countries switched the mode of administration from PBA to CBA, but there were still some countries that remained with the PBA. In PISA 2022, some of those countries also switched to CBA.

As the trend contents need to remain identical across cycles, the transfer of trend contents from PBA to CBA, i.e., from Word to XLIFF, was centrally managed, as it was in PISA 2015 and PISA 2018. To do this operation, a semi-automated process (different from the more manual process applied in 2018) was adopted. National centres were then asked to review their transferred units using the preview widget on the PISA portal and report any transfer error or residual issues identified in the trend materials using change request forms (in Excel format). Approved changes were then centrally implemented by the contractors.

The workflow of the trend transfer process is shown in Figure 7.2. It details the two parallel workflows that have been developed to transfer the content of the Trend PBA units into the new CBA format. First the PBA materials were extracted from Word and aligned to produce a Translation Memory (TM). Then, the new content that was specific to CBA environment, like specific instructions such as “Click on”, or “Select”, were translated so that these could already be used to pre-translate the CBA xliffs. Once this pre-translation phase was completed, Quality Assurance checks were performed and translated segments were locked in the OmegaT projects. These transferred materials were then uploaded to the PISA portal for the Countries/economies to review. Any residual issue was then documented by the Countries/economies and corrected centrally. The Countries/economies did not have editing rights to trend content at any stage of the process. This approach prevented unnecessary, undocumented, or unverified changes in the trend materials, and thus will allow both more reliable comparability across cycles, and a detailed record of all changes made in trend materials.

Figure 7.2. Trend Transfer process diagram



Questionnaire adaptation negotiation

Questionnaire verification before the Field Trial aims to ensure cross-linguistic equivalence of the national versions of the data collection instruments. This process began with the negotiation of national adaptations documented in the Questionnaire Adaptation Spreadsheet, referred to as QAS in this report.

In the questionnaires, national adaptations are defined as intentional deviations from the source, aiming to reflect the national context and to keep the comparability on the international level at the same time. A set of these national adaptations was compulsory, such as country-specific response options in a question that asks about education levels, types of school, or language spoken at home. Beyond these "forced adaptations", countries could propose requests for additional adaptations in the QAS.

Countries proposed their adaptations to new items in the QAS and provided a back translation in English and a justification for the adaptation, as needed. Based on the back translation and the justification, the questionnaire team either agreed to the proposed changes, or asked the National Centre to further adjust the translation to correspond to the source and ensure across-country comparability. This dialogue between the National Centre and the contractors took place in the QAS until an agreement was reached. Then the country-specific "national source" was created by the questionnaire team.

The National Centre implemented the agreed adaptations in their national versions. CBA countries encoded the adaptation directly in the QAT.

After having tested the different scenarios (rules and filters) advised by Core A (ETS Data Management), countries uploaded the QAS document detailing the negotiation and released the national questionnaires for the next step in the workflow, i.e., verification.

For the first time in PISA 2022 the questionnaire verification was aligned with the cognitive materials in terms of technology, which meant using OmegaT for both. When the negotiation of national adaptations was completed, a national source was created on the QAT. The national source was then exported from QAT in XLIFF format for the use in OmegaT. Trend items were centrally populated in OmegaT and locked for editing. The Countries/economies had the possibility to request the changes to trend items within the QAS. These change requests were then negotiated with the Questionnaire Content Team and if agreed, implemented by the verifier during the verification step.

For Countries/economies switching from PBA to CBA, translation memories were created from the PISA 2018 Word files of the questionnaires and transferred on the QAT. The translation of trend questions was thus pre-populated in OmegaT, and all players were instructed to align the translation of the new questions to the existing trend translations.

For PBA versions, the Countries/economies were responsible for maintaining their trend translations.

International verification of the national versions

As in previous PISA survey administrations an independent team of expert verifiers were appointed and trained by the international contractors to verify each national version against the English and/or French source versions to ensure high-quality standards and assessment materials and contextual questionnaires.

New CBA test units

Of the 88 Countries/economies participating in the PISA 2022 Field Trial, 5 participated in the paper-based assessment (PBA). The remaining 83 Countries/economies participated in the CB assessment.

CBA units were translated and verified using XLIFF files on OmegaT. The files were exchanged, previewed and archived on the PISA portal, a web-based platform that allows the files to travel through a predefined workflow.

To perform the verification task, the verifiers were instructed to compare the translated segments to the source one by one in OmegaT, while consulting previews on the portal and checking item-specific guidelines and comments from the national centres in the TAS. Where corrections were needed, the verifiers implemented them in OmegaT and documented their interventions in the TAS, using a predefined drop-down menu to assign the change to the appropriate intervention category.

Once a domain was verified, reviewed and finalised on the portal, the translation referee was able to download the TAS annotated by the verifier. The referee would then go through each verifier and country comment, and label as “requires follow-up” any crucial issues that could potentially affect equivalence or item functioning.

Changes labelled as “requires follow-up” were negotiated between the referee and the national centre. The national centre then uploaded revised OmegaT packages and TAS on the portal 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 international contractors 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 test adaptation spreadsheet. Figure 7.3 shows a sample test adaptation spreadsheet from the PISA 2022 Field Trial. The spreadsheet functions as:

- an aid to translators, reconcilers, and verifiers through the increasing use of item-specific translation/adaptation guidelines,
- a centralised record of national adaptations, of 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 7.3. Sample of a TAS from the PISA 2022 Field Trial

ENGLISH SOURCE VERSION	ITEM-SPECIFIC TRANSLATION/ADAPTATION GUIDELINE	PARTICIPANT COMMENT (ADAPTATIONS, DOUBTS, DIFFICULTIES)	INTERVENTION CATEGORY	VERIFIER COMMENT	Corrected?	TRANSLATION REFEREE COMMENT	CORRECTION STATUS	PARTICIPANT POST-VERIF COMMENT	VERIFIER FINAL CHECK	VERIFIER FINAL CHECK COMMENT
She is correct, because the height of a medium box is 2/3 the height of a large box.	Pattern: response options start in the same way. If possible, reflect the pattern in the target.		OK	T&A guideline followed.						
She is correct, because 3 medium boxes can always be fit into the same space as 2 large boxes.										
She is not correct, because none of the interior storage dimensions of truck A are multiples of 0.75, which is the height of a large box.	Please do not use the translation of the word 'area' in reference to the compartment because this unit is all about volume.		Inconsistency	"box" translated inconsistently within item. T&A guideline followed.	Yes	Please keep the verifier correction	REQUIRES FOLLOW-UP	ok	OK	

Cognitive trend units

For cognitive trend units, i.e., units that the Country/economy has administered in one of the previous cycles, it is essential that the unit is administered in the exact same form to be able measure trends in time. For this reason, centralized trend management was deployed. The Countries/economies did not have editing access to trend units, i.e., units that the Country/economy had administered in one of the previous cycles, at any point of the translation, adaptation, and verification workflow. They were given the opportunity to request changes to trend units, if for example a residual linguistic error or outdated adaptation was identified. The Countries/economies documented these requests with a justification for change in a change request form (Excel file). If the translation referee and the verifier agreed that a change is indeed acceptable, it was implemented by the verifier.

The verification workflow for the trend units is shown in

Translation/Initial review: Country used the preview widget to review the trend content transferred; and documented any errors found or changes they would have liked to make likewise in the Change Request Form.

Negotiation: Country negotiated change requests with Translation Referee who either accepted or rejected each request made by country based on the information provided.

Referee Review: Referee approved or rejected country's change requests.

Verification: Verifier evaluated linguistic correctness of change requests from country and implemented change requests accepted by Referee in the XLIFF.

Verification Review: cApStAn staff reviewed the verifier's feedback.

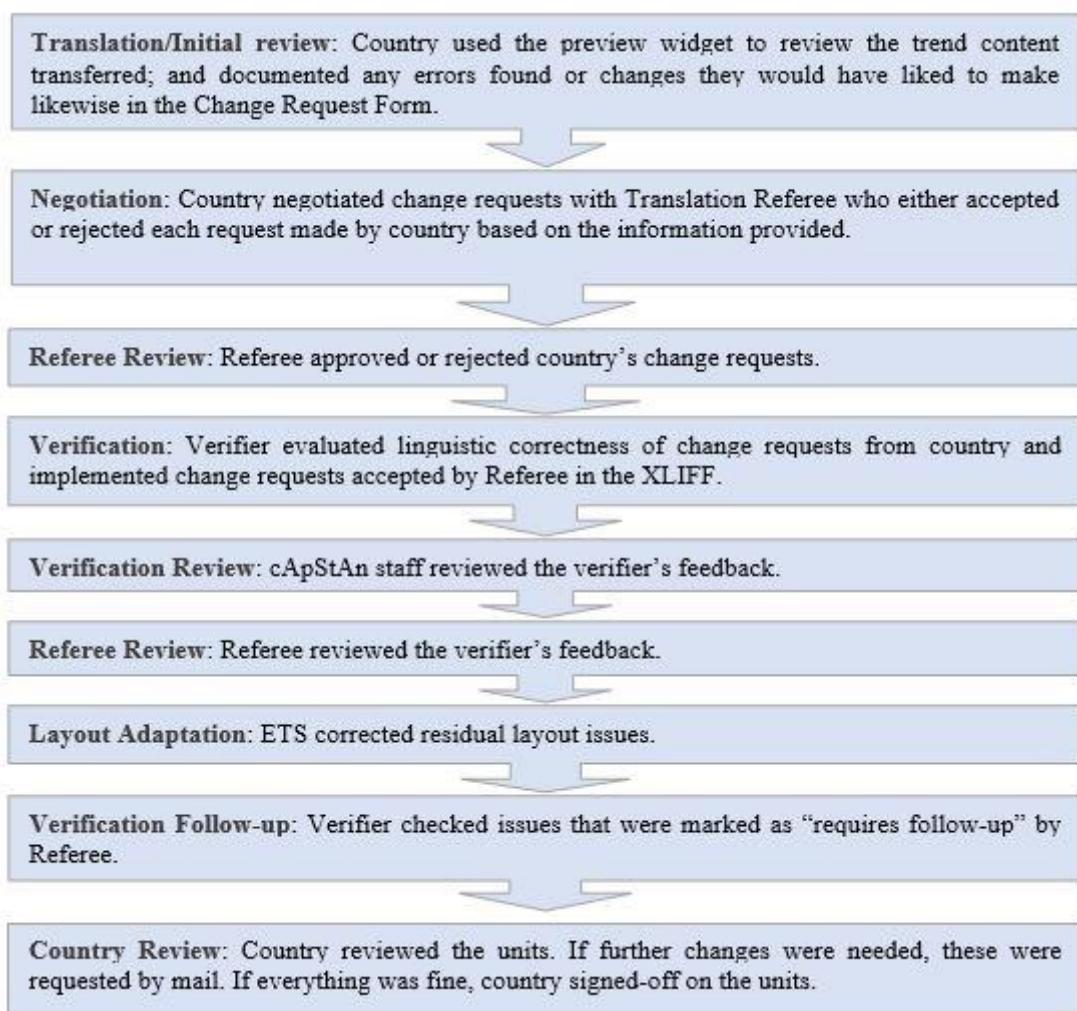
Referee Review: Referee reviewed the verifier's feedback.

Layout Adaptation: ETS corrected residual layout issues.

Verification Follow-up: Verifier checked issues that were marked as "requires follow-up" by Referee.

Country Review: Country reviewed the units. If further changes were needed, these were requested by mail. If everything was fine, country signed-off on the units.

Figure 7.4. Verification workflow of trend items



For trend items there was no difference between adapted and translated versions as regards the Change Request Form and the overall procedure.

As a National Centre reviewed the Trend items, it reported any linguistic or content-related request for modification and then submits the annotated Trend Change Request Form to the Translation Referee for approval. All errors related to the trend transfer procedure, that were thus not changes versus trend content, were automatically approved. For any requests that would mean a real content related change versus trend the referee's role was to evaluate whether the requested changes were legitimate or not and could have an impact on the trend data collection. The result of this arbitration process was a Change Request Form where the countries' requests were either approved or rejected by the referee.

The following type of change requests were generally accepted:

- requests to correct outright errors, such as typos, blatant grammar issues, mistranslations,
- requests to correct outdated adaptations, e.g., change of currency,
- changes to harmonise form of address (informal/formal 'you') across materials coming from different cycles,
- requests to harmonise spelling following a spelling reform,
- requests to harmonise decimal and thousand separators across items, and

- changes to improve wording or to correct errors in an item that has not performed well and showed Differential Item Functioning in previous cycles

The following type of requests were generally rejected:

- preferential changes, improved wording when there are no statistics showing that the item has not performed well in the past,
- punctuation issues,
- capitalization issues (unless outright errors),
- changes that would be against the translation and adaptation guidelines, and
- changes to bring the target version closer to one source version while it already corresponds to the other source version (i.e., changes introducing expressions idiosyncratic to English when a version has been translated from French).

The verifiers' brief for trend verification was to implement the changes that had been requested by the national centre and approved by the referee, if linguistically appropriate – or if not appropriate, suggest a revised wording.

Once changes were verified and implemented in the XLIFF files, verifiers double-checked on the preview that everything appeared correctly in the preview. In principle, no other changes were allowed unless typos or blatant errors were discovered. If the verifier spotted other mistakes that could affect the trend nature of the items, the referee's judgment was called for. At the end of the process, the verifier uploaded the updated XLIFF and Change Request Form files on the portal. The verification step was followed by an internal review by cApStAn. At the end of the process, the files were uploaded on the portal and pushed to Layout Adaptation step.

After the layout adaptation step, the files were pushed to cApStAn for a final check. The main aim of this step was to double-check that all layout issues pointed out during the verification and the review processes had been addressed and to correct any residual issues.

At this stage, the procedure therefore consisted in:

- double-checking that the most important errata (including latest errata released after verification and review) had been implemented
- making sure that the layout issues had been addressed
- addressing any residual issues

If residual layout issues were found, the relevant files were sent back to ETS for further correction and another check was performed thereafter. At the end of the process, all the files were uploaded on the portal for the national final check and sign-off.

After verification follow-up the countries had a last opportunity to check that all new translations and relevant accepted changes had been implemented correctly and that any residual layout issues, whether raised by the countries themselves or by the verifiers, had been addressed. If errors were still encountered this needed to be commented in the Change Request Form.

The final sign-off from the National Centre ended the trend verification procedure.

Questionnaires

The successful administration of questionnaires in large multinational, multicultural and multilingual surveys depend heavily on their correct adaptation to the national context. The comparability of the data is guaranteed by “asking the same question” in all the Countries/economies and in all the languages, and to this end, the first task of the PISA Countries/economies was the negotiation of the adaptations, before the translation started.

Questionnaires were submitted for verification together with an agreed QAS. The first purpose of the questionnaire adaptation spreadsheet was to document all content-related or ‘structural’ deviations from the international reference versions. Such national adaptations were subject to approval by the questionnaire team before the material was submitted for verification. Subsequently, the spreadsheet served the same objectives and followed the same logic as the test adaptation spreadsheet for test units (see above). Figure 7.5 shows a sample questionnaire adaptation spreadsheet from the PISA 2022 Field Trial.

Figure 7.5. Sample of a QAS from the PISA 2022 Field Trial

The verifiers' brief was to check whether the target questionnaires are linguistically correct and faithful to either the source version (when no adaptation is made) or to the approved English translation of the national version (when an adaptation is made). In light of this, verifiers were instructed:

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

For the paper-based questionnaires (Student and School questionnaires for countries administering paper-based assessment, Parent Questionnaire for all Countries/economies taking this option), verifier interventions were entered in the questionnaires using the track changes mode, while verifier comments were entered in the verifier columns of the questionnaire adaptation spreadsheet.

For CB questionnaires the verifier applied necessary interventions on OmegaT and documented the rationale for the change in the QAS.

When the verification was completed, the Questionnaire Content Team reviewed the verification feedback and labelled as “requires follow-up” important issues that could potentially affect cross-country comparability. The files were sent back to the country/participant for their review before going through the last passage of Final Check.

The translations of the Global Module Crisis module were produced following a different workflow. cApStAn produced the translated materials through the double translation and reconciliation model. Countries/economies reviewed the translations and requested changes or national adaptations through the QAS. The Questionnaire Content Team assessed the requests and indicated if they were approved or not. The files were then transferred to the verifier who implemented the agreed corrections/updates. There were no special procedures for the verification of the questionnaires adapted from the source versions, from the common base versions or from borrowed versions, since differences in education systems mean that these are very extensively adapted even when sharing a common language. Nevertheless, English and French versions benefited from a co-ordination process similar to the one implemented for test materials. A list of “tips” for verification of questionnaires, including spelling, possibly recurring adaptation issues, and especially errata (errors identified in the source version after release to the Country/economy) and “quasi-errata” (suggestions for improving the source) was maintained, built up, and used in each successive verification.

As in previous cycles of PISA, there was also an increased effort to harmonise the verification feedback for different language versions of questionnaires used in the same country (e.g., German, French and Italian for Switzerland, or the four language versions for Spain). Such versions are by necessity entrusted to different verifiers, but when possible, cApStAn's verification reviewers aimed to review and deliver such versions together, striving to harmonise verification interventions on adaptation issues common to the different language versions.

Adapted versions

Whenever a country adapted their national version from the English or French source, a common base version, or verified version from the same language borrowed from another country, this was considered an adapted version. The resulting national version was verified using a special procedure for these versions. There were in total 50 CBA adapted versions that were verified using this process.

The essential difference between the “full” verification of translated national versions and the “focused” verification of adapted versions is that in the latter, the verification concentrates on the changes made by the country versus the source, common base or borrowed version. Automatically created difference reports were used to identify all such changes in a reliable way.

Paper-based test units and booklet shell

Since no new paper-based units were developed for PISA 2022, PBA Countries/economies that had participated in previous cycles did not have anything new that required translation or adaptation. For these Countries/economies, the units only went through the centralised change management process whereby the Country/economy had the opportunity to request corrections to errors, and these – when accepted by the translation referee – were then implemented centrally by the verifiers.

Paper-based countries that were new in PISA 2022 or that had not participated in one or more of the relevant cycles had to translate or adapt units they had not administered before. These were verified following the same process as described above for CB materials. The only essential difference was that the verifiers implemented the changes in the Main Survey Word files using the “track changes” functionality, rather than in OmegaT. The test adaptation spreadsheet was used the same way as in the CB verification.

Coding guides

In PISA 2022, the coding guides were verified separately from the test items, and at a later time. This was necessary since many additions and improvements were made to the master versions after the coder training meetings, long after preliminary versions of the guides had been made available to Countries/economies. As in PISA 2015 and PISA 2018, the scoring sections were not made available for translation at the time of the unit dispatch. There was one coding guide per trend domain (mathematics, science and reading). For CBA Countries/economies, there was, in addition, one coding guide for New Math, and for those Countries/economies that opted for Financial Literacy and/or Creative Thinking, there were separate coding guides for these domains.

As opposed to the previous cycles, in this cycle the new coding guides were verified using OmegaT. To be able to use the latest version of the translation memories of the cognitive units, the workflows for the coding guides were created only after the cognitive materials were verified. The overall verification procedure was the same as with the cognitive units. The verifiers made corrections as needed in OmegaT, documenting their interventions in the coding guide adaptation spreadsheet (CAS), including selection of the appropriate intervention category using a drop-down menu. However, there was a significant difference between the verification of the cognitive units and the verification of the coding guides: The translated files for the coding guides were in Main Survey Word format and therefore layout issues had to be corrected manually after the verification process had been completed.

The New Math coding guide went through a full verification in the Field Trial. For the Main Survey, central revisions to reflect updates to the source were made by the Countries/economies in OmegaT together with additional changes which were deemed necessary to correct errors. The verifiers were asked to review both the updates and the edits.

To accommodate the changes to the Creative Thinking coding guides after the Field Trial International Coder Training, the OECD and contractors determined it was important to devote more time to produce updated source versions. Due to time constraints, there was no verification of the Field Trial Creative Thinking coding guides. Instead, a full verification was implemented for the Main Survey.

The Creative Thinking master coding guide was updated after the Field Trial, and the Countries/economies were asked to reflect these updates in their translations. They did this in a newly generated OmegaT project where the translation memories from the revised Main Survey units as well as the translation memories from the Field Trial Creative Thinking coding guides were included. While implementing these central updates in their translations, the Countries/economies also had the opportunity to correct residual errors detected during their review of their Field Trial data.

For Countries/economies that had participated in previous cycles, trend coding guides underwent a similar controlled change request process as for the test units.

Outcomes of the Field Trial verification

The TAS and the QAS in Excel format were used to document the verification of test units and the questionnaires. For each issue they encountered, verifiers were required to choose from a drop-down list of 14 intervention categories and then explain the details of the issue and of their intervention in a comment.

The predefined intervention categories in the drop-down menus of the TAS and QAS are linked to formulae, which generate statistics on the number and types of verifier interventions in test units, both per language version and per unit. The data is available in detailed form in Appendices 4-8 of this chapter (in Excel format). In this section, some of the data will be presented, together with some figures and graphs.

For reasons of comparability, the data of the translated versions are shown separately from the data of the versions that were adapted from the French or English source versions or from the Chinese or Spanish base version, or from a verified national version of another country. For these adapted versions, the process was different as it was a focused verification of national adaptations proposed by the national centre, rather than a full sentence-by-sentence verification. The results are not comparable with the translated versions where the whole translation was verified sentence by sentence.

The statistics in this section cover national versions of New Mathematics units and Creative Thinking units. The list of language versions is not identical between the two domains for two reasons: some National Centres opted out of the Creative Thinking innovative domain, and for some other countries the Translation Plan was different depending on the domain. Also, some countries opted for a hybrid plan; for example, for the New Mathematics units Bosnia and Herzegovina, Serbia and Montenegro translated a third (one batch) of the units each and adapted the other two thirds, so in the statistics they appear in both tables and graphs.

For each national version included in the analysis, the formulas embedded in each of the TAS produced the following figures:

- the total number of verifier interventions in the 61 New Mathematics units across the 113 language versions;
- the total number of verifier interventions per intervention category in these units; and
- the total number of verifier interventions “requiring follow-up” and related percentage.

In addition, for each unit, data was extracted to obtain:

- the total number of interventions per intervention category (in translated and adapted versions);
- the total number of interventions “requiring follow-up”; and
- the percentage of each type of intervention category vs. the total number of issues reported.

While figures per national version can be informative, they need to be interpreted with care. An illustrative sample of possible scenarios is presented below.

Two versions are of the same generally acceptable quality. One is verified by a strict verifier who extensively comments on even minor errors; another is verified by a more pragmatic verifier who documents only major issues. The statistics might show a great number of interventions in the first version, and considerably less in the other. This difference in verification styles should, however, show in the percentage of interventions “requiring follow-up”, which should be lower than average in the version verified by the “strict” verifier.

One verifier may have reported an “Inconsistency” issue in the TAS every single time the issue appeared. Another verifier may have chosen to report such cases only once, with the note “Corrected throughout the units without further comments” in the verifier comment on the first occurrence. Similarly, one verifier may have reported a recurring issue (e.g., a repeated ‘mistranslation’) each time it occurs, while another verifier might cover that with one generic comment.

Recurring issues, such as missed harmonisation of repeated instructions or inconsistency in form of address, generally labelled as “Inconsistency”. If the number of such interventions is very high in a version this may be due to the fact that trend translations were not considered when translating or adapting the new units.

There may be several separate issues in one sentence/paragraph that the verifier has documented in the same row in the TAS. As only one category can be selected per row, it would be selected according to the most severe issue.

In adapted versions the verifiers are mainly focusing on national adaptations vs. the base and correct implementation of the errata. This explains the fact that these two categories appear to be much higher in adapted versions versus translated versions.

While looking at the total number of interventions does give some indication of the translation quality of the national version, it does not take into account the severity of the issues discovered by the verifier. It makes more sense to look at several combined factors that may serve as indicators for translation quality. One should examine the total number of changes labelled by the Translation Referee as ‘requiring follow-up’ and the number of issues in the more ‘severe’ intervention categories – mistranslation, adaptation issue, matches & patterns, and guideline not followed.

New cognitive items: translated versions

Even if most of the verifiers rated the translations as very good or good, the verifier interventions were key to maintain the linguistic equivalence to source and correct any residual language issues.

In the translated versions of the New Mathematics units, the categories which revealed the most verification interventions were:

Minor linguistic issue – this category is used for typos or other linguistic defect such as spelling, grammar, capitalization, punctuation, etc., that does not significantly affect comprehension or equivalence. Correcting such errors is usually not controversial, and in the Mathematics units 25% (see Figure 7.6) of the verifier’s interventions fall into this category.

Inconsistency – typically used for interventions when an element across units (e.g., an instruction or prompt) is inconsistently translated, and it is not intentional or documented as an adaptation. The verifiers’ corrections show 18% in this category, as shown in Figure 7.6.

Grammar or syntax – this category was used to document 13% of the verifiers' interventions (see Figure 7.6). It is used for corrections of grammar mistakes that could affect comprehension or equivalence, e.g., wrong subject-verb agreement, wrong case (inflected languages), wrong verb form, or syntax-related deviation from the source and was used in 13% of the interventions, as shown in Figure 7.6.

The low percentages of corrections of severe translation issues such as mistranslation (6%) or adaptation issues (4%) shows the good quality of the translation (Figure 7.7). No corrections of the matches and patterns were recorded in these units. This deviation from the source of is typically more frequent in Reading literacy units' literal matches (repetition of the same word or phrase) or a synonymous match (use of a synonym or paraphrase) or patterns in multiple choice items (e.g., all but one option start with the same word, proportional length of responses options) need to be reflected in the target version for valid data measurement and comparison.

Figure 7.6. Distribution by category of verifier interventions in New Mathematics units (translated versions)

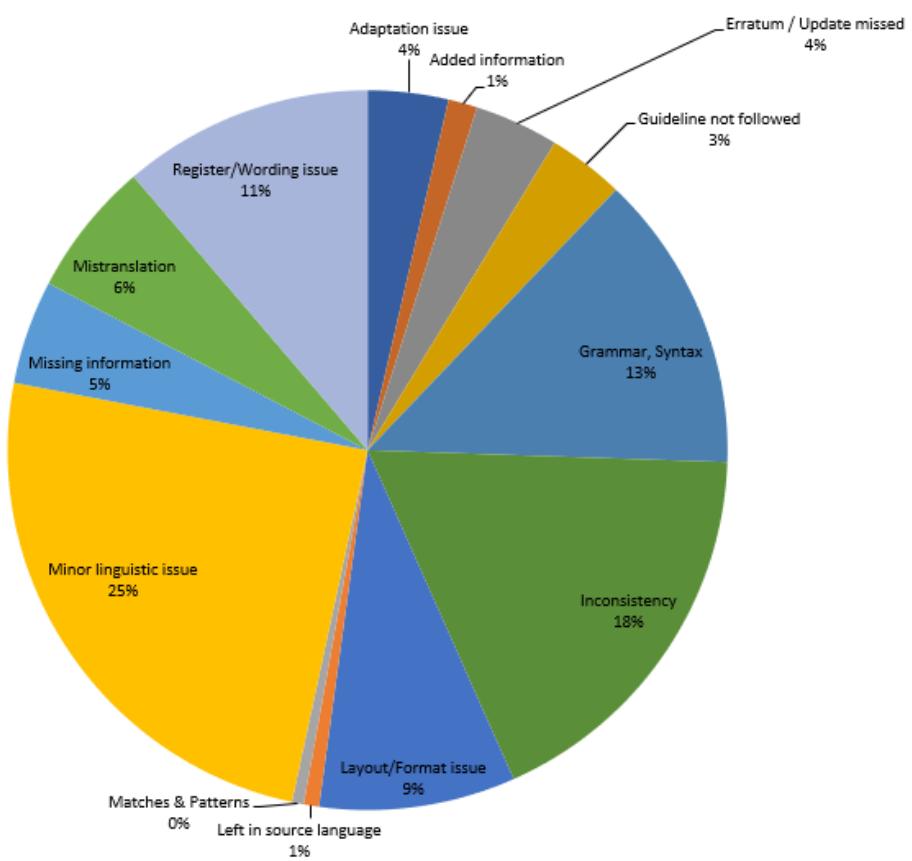
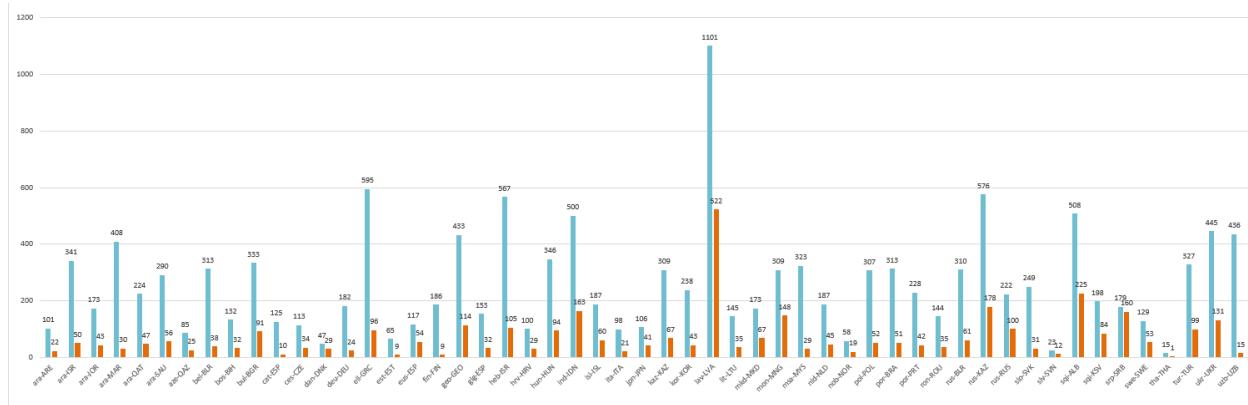


Figure 7.7. Number of issues per national version in New Mathematics units (translated versions)



The outcome of the verification of the Creative Thinking units is similar, with 27% of interventions were for corrections of inconsistent translation and 18% of corrections of minor linguistic issues, as shown in Figure 7.8. The number of issues per national version can be also found in Figure 7.9.

Figure 7.8. Distribution by category of verifier interventions in Creative Thinking units (translated versions)

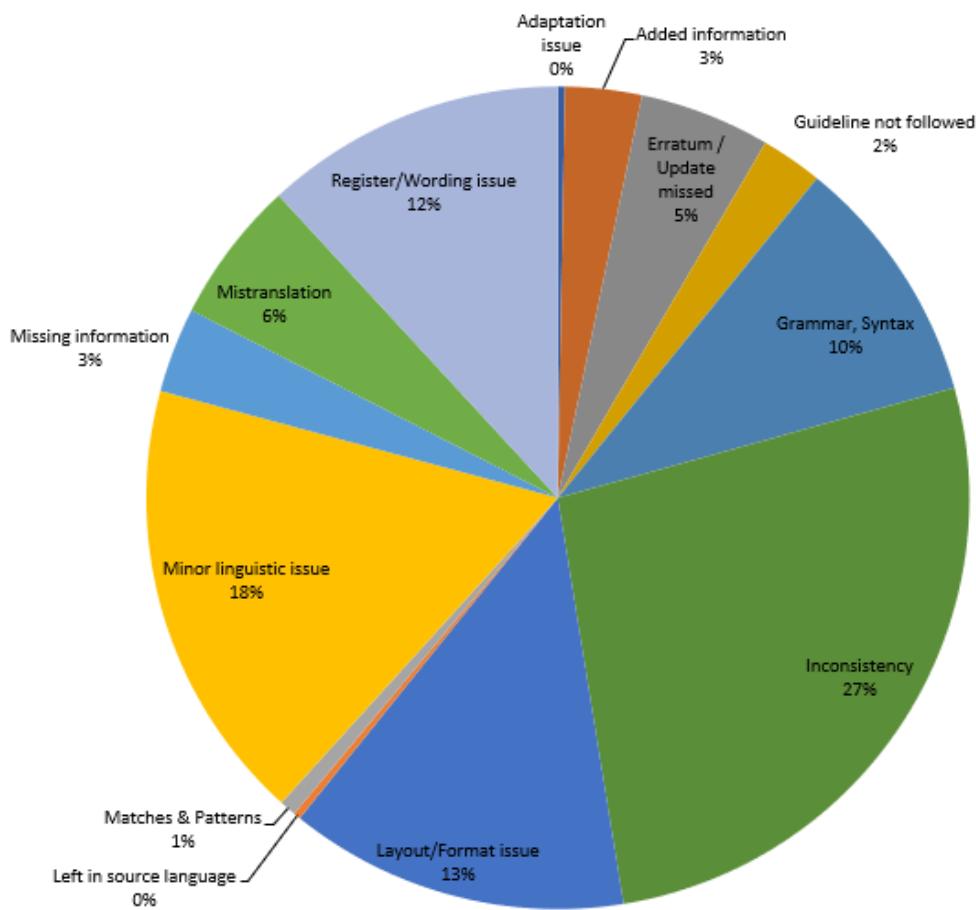
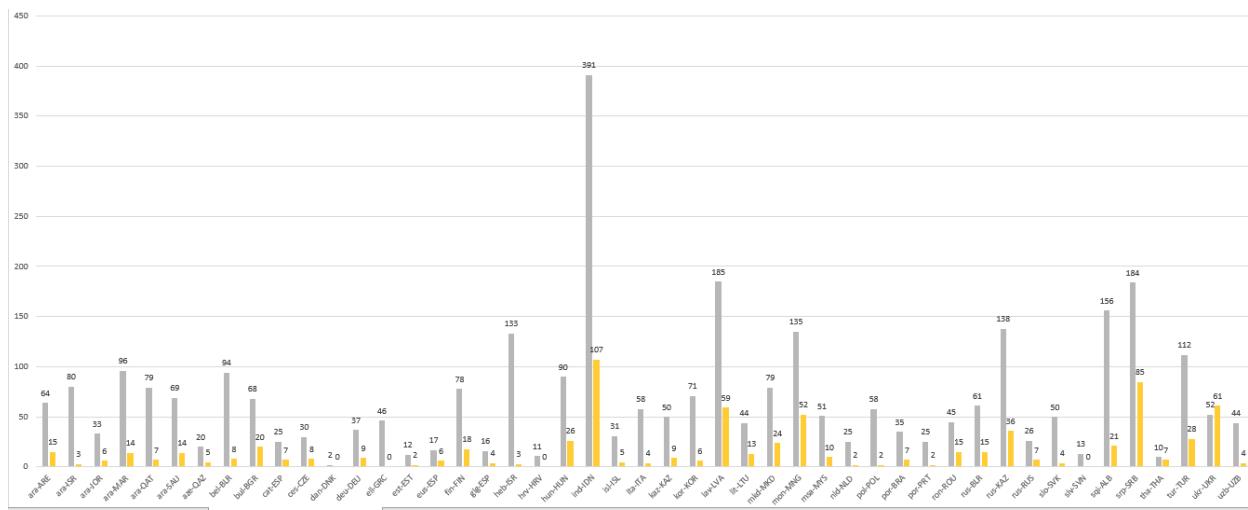


Figure 7.9. Number of issues per national version in Creative Thinking units (translated versions)



New cognitive items: adapted versions

For the versions adapted from the English or French master version, from the Chinese or Spanish common reference version, or from a borrowed verified national version, the issues identified by verifiers mostly belonged to the following types:

Adaptation issue – As shown in Figure 7.10, in 17% of the verifiers' interventions, required adaptation was missed, materials were not adapted at all or poorly adapted; adaptations was not correctly or consistently implemented. For example, the adaptation documented in the TAS was not implemented as described in the XLIFF file, or implemented only in some occurrences; adaptation or change proposed by national centre was not acceptable (e.g., it added information not present in the source or made the national version easier or more difficult). Typical examples of adaptation issues in adapted versions are: missed adaptation of spelling and typographic conventions (e.g., UK to US English spelling, date formatting, decimal and thousands separators), fictitious character names not adapted to local context, etc.

Inconsistency – similar to the translated version, 19% of the corrections fall into this category (See Figure 7.10).

Minor linguistic issues were corrected in 14% of the interventions, errata were corrected in 12% of the interventions and layout or formatting such as emphasis (bold, italics, underline) was adjusted in 10% of the interventions in the adapted versions of the New Mathematics units, as shown in Figure 7.10.

Figure 7.10. Distribution by category of verifier interventions in New Mathematics units (adapted versions)

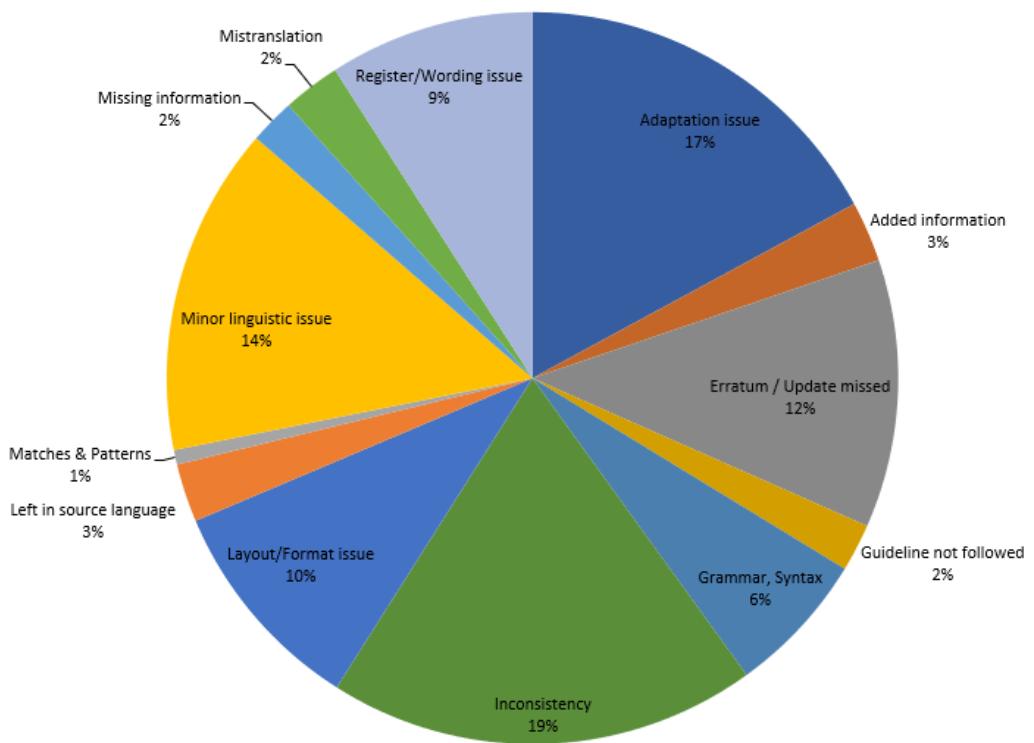
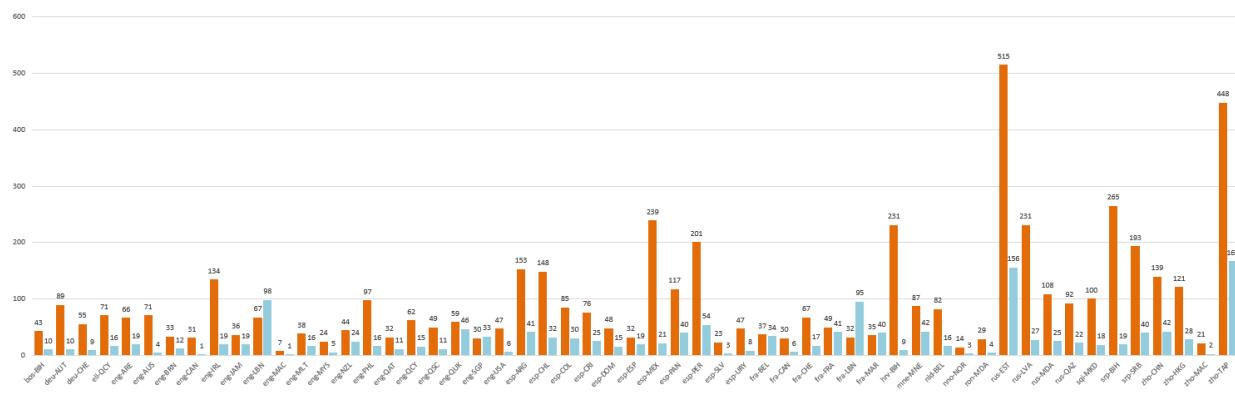


Figure 7.11. Number of issues per national version in New Mathematics units (adapted versions)



For Creative Thinking, inconsistencies were harmonised in 21% of the verifiers' interventions, errata were corrected by the verifiers in 18% of their interventions, and register, wording and minor linguistic issues were corrected in 12% of the recorded interventions, as per Figure 7.12. In addition, Figure 7.13 presents a breakdown of issues per national version.

Figure 7.12. Distribution by category of verifier interventions in Creative Thinking units (adapted versions)

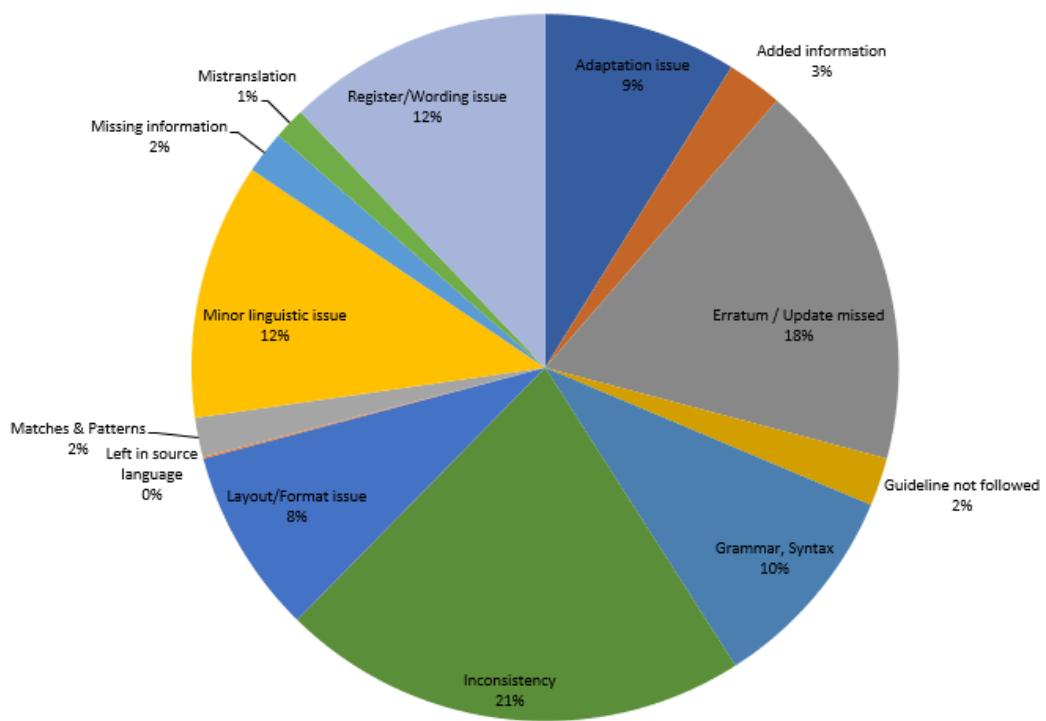
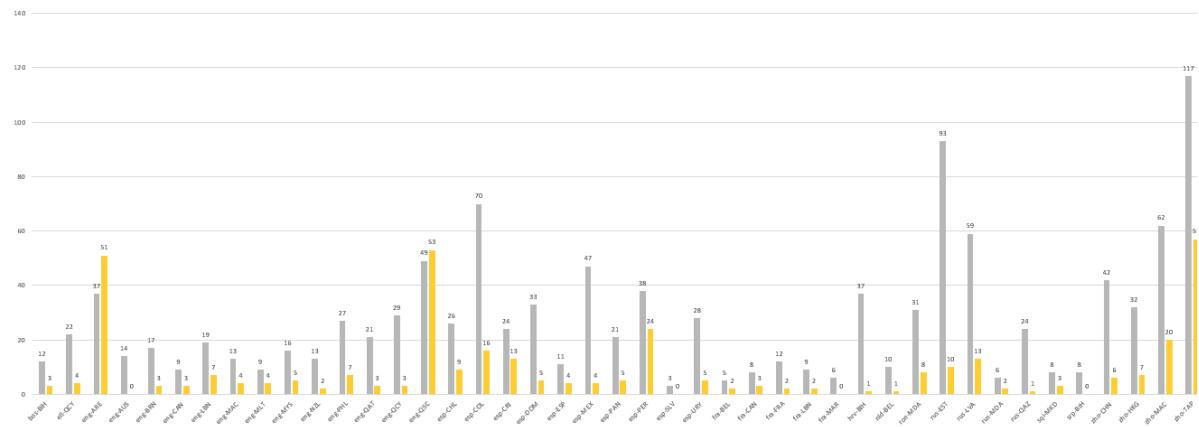


Figure 7.13. Number of issues per national version in Creative Thinking units (adapted versions)



Main survey verification

Cognitive units

As in PISA 2018, no major changes were made in the master versions after the Field Trial (apart from entire units or items being dropped) in PISA 2022. The changes that Countries/economies requested to their Field Trial instruments, for example based on poor performance or differential item functioning in the Field Trial, or the detection of residual “outright errors” needed to be verified and centrally implemented together with the implementation of the FT-to-MS errata. These errata included errata discovered after the

last release of the Field Trial errata document and central Field Trial to Main Survey updates. This process was similar to the centralised change management used to control changes in trend: Countries/economies requested changes, and the verifiers implemented centrally those changes that were approved by the translation referee. The Countries/economies did not have editing access to their units or questionnaires at this stage.

The trend items and new items followed the same workflow although the process for New and Trend materials was slightly different.

The Main Survey preparation started after the Main Survey item selection was confirmed. The release of the Main Survey workflows was linked to the data release, so the timeline strongly depended on the compliance of the National Centres in submitting their Field Trial data.

During the Main Survey review countries were asked to carefully review any items that did not perform well in the Field Trial and try to identify whether this country-item interaction was language-driven. Such items were highlighted in red in the "item feedback form" (IFF) in Excel format and indicated by a "YES" in the "Flagged Item?" column.

In case the National Centre spotted residual errors, they had the opportunity to request changes to the translation. Changes had to be requested in the IFF where countries were asked to enter: a short description of the error, the location of the error (e.g., segment number), the English or French source for that segment, the original Field Trial wording, a back-translation of the original Field Trial wording and the proposed corrected Main Survey wording.

There was one item feedback form for all cognitive items with a separate tab for each domain (New Maths, Trend Reading, Trend Science, Trend Math, XYZ and Trend-New Financial Literacy, see Table 7.6 for an example). The IFF also included 2 Instructions tabs describing in detail the process for the new and trend materials. At verification 2 additional columns indicating the dropped items and the Main Survey errata were added.

There was one single workflow for all the Core instruments for the Field Trial to Main Survey verification which included two Referee review steps, one *before* the verification which was used to review national centre requests for changes in the trend materials, and one *after* verification, to flag any major issue, as usual.

All national centre requests were reviewed by the verifier, who double-checked (i) whether it was an outright error or a preferential change, and (ii) whether the proposed Main Survey wording was still equivalent to the source and linguistically correct. As a general principle, for the trend materials the principle of identicalness of trends was applied and any preferential change or change to a non-flagged item was generally not agreed by the Referee and therefore not implemented by the verifier. Agreed corrections were corrected in the XLIFF files by the verifier. Additionally, verifiers were responsible for implementing all Field Trial to Main Survey errata, that is errata which were discovered between the Field Trial and the Main Survey.

Countries did not have access to the XLIFF files at any point of the process; all changes were implemented centrally by cApStAn verifiers. Countries could nevertheless consult the unit previews and DIF reports at different stages of the process, to make sure their requested changes and the Field Trial to Main Survey errata were correctly implemented during verification.

The general guideline of correcting only outright errors (and, more generally, the concept of "outright error") was not understood and accepted the same way by all countries. Some only requested a few justified changes, others called for a more extensive revision of the units (e.g., Kazakhstan, Mongolia).

Questionnaires

As in Cognitive items, no content-related changes were made in Questionnaires items that made it to the Main Survey. The changes in the questionnaires before the Main Survey were mainly structural. Full questions and response options within items were omitted, the order of questions was changed. Finally, a couple of updates in two questions and an introductory part were implemented centrally by ETS.

The structural changes were implemented by ETS in the QAT for the countries that administered the questionnaires on computer. For the paper-based questionnaires, the National Centres reflected these changes in their materials in Word format before generating the final questionnaires in PDF format.

The procedure was similar to the procedure for the cognitive units. The few content changes such as the addition of the consistency checks for scale questions were considered as errata and were added to the necessary update in the year of administration in SC002 and the Field Trial to Main Survey errata. A tab for the documentation of these updates, the Main Survey Questionnaire Change Request Form was added in the QAS (Figure 7.14). The QAS also contained the locked Field Trial QAS tabs for reference, well as a tab with an example of correctly documented Main Survey Changes.

In the Main Survey Questionnaire Request Form countries could also request other updates due to objective major modifications (e.g., changes in the school programs on national level), or ask for correction of errors in items showing strange behaviour in the Field Trial data. They were advised against any changes in items that worked well in the Field Trial.

The Questionnaire Team at ETS reviewed the documented updates and possible requests for corrections of errors and recommended their implementation when applicable.

At verification stage, the verifier checked the linguistic correctness of the update in the target language and implemented centrally to the questionnaires in XLIFF the agreed changes. In the step after this implementation, countries could review it in the QAT, and reported in the QAS if any residual issues needed to be addressed.

The same procedure was followed for the PBA materials, with the difference that the National Centres reflected the recommended updates and agreed corrections in the questionnaires administered on paper.

Figure 7.14. Main Survey Questionnaire Change Request Form in the QAS

PISA 2022: MS Questionnaires Change Request Form					Step 6.1 (MS) - Identify Changes		Step 6.2 (MS) - Approve Changes		ETS Questionnaire Content		
PISA Centre			6.1 Fill in these 5 columns for all corrections		6.2 Fill in these 3 columns for all translation corrections			6.2 Comments on correction requests		6.2 Approval Status	
Questionnaire (Select from the dropdown (e.g., ST001))	Question ID (e.g., ST001Q001TA01)	Item ID (e.g., ST001Q001TA01)	Type of Correction (Select from the dropdown menu)	Description of correction	FT Translation, in National language (Full segment)	Requested Translation for in National language (full segment)	Back-translation of requested MS Translation, in ENGLISH	6.2 Comments on correction requests	6.2 Approval Status		
TCQ	TC258	TC258E01	ERRATUM	Addition of Consistency Check for Scale Questions to avoid items being listed as unanswered in the Questionnaire. PISA Centres should translate the text in Column H into their national language, and include the updated translation in	N/A	"0"이라고 응답하기 위해서는 는 글의 슬라이더(조절 단추)를 "0"의 위치로 이동하세요.	To enter a response of "0" (zero) for a question, please move the slider to the "0" position on the scale.		AGREED		
TCQ	TC259	TC259E01	ERRATUM	Addition of Consistency Check for Scale Questions to avoid items being listed as unanswered in the Questionnaire. PISA Centres should translate the text in Column H into their national language, and include the updated translation in	N/A	"0"이라고 응답하기 위해서는 는 글의 슬라이더(조절 단추)를 "0"의 위치로 이동하세요.	To enter a response of "0" (zero) for a question, please move the slider to the "0" position on the scale.		AGREED		
TCQ	TC261	Description	DELETION	Description removed from this item. Please review item screen to ensure that the description has been properly deleted.	Please consider your employment status at this school and for all of your teaching employments together.	Deleted	N/A		AGREED		
TCQ	TC261	Instruction	ERRATUM	Incorrect instruction text. Instruction changed to (Please select one response.) for this item. PISA Centres should translate the text in Column H into their national	(Please select one response in each row.) (가 항목에서 하나를 선택하십시오.)	(하나를 선택하십시오.)	(Please select one response.)				

Coding Guides

The coding guides for the new cognitive items were translated and verified in XLIFF format, therefore the Main Survey updates and corrections of the errata followed the same procedure as the instruments.

For the Main Survey, the countries were asked to produce Main Survey versions of their trend coding guides starting from their final Field Trial versions, reflecting all applicable revisions made in the master versions. Separate Main Survey master versions were produced for PB and CB countries. The Field Trial to Main Survey revisions that countries were asked to reflect were of the following types:

- Removing scoring sections of items that did not make it to the MS
- Making edits in the cover, footers and introduction
- Reflecting Field Trial to Main Survey revisions the test developers made in the scoring sections, e.g., modifications in the scoring instructions or addition/removal of sample responses

Master versions with tracked changes were released to countries and they were asked to reflect all the Field Trial to Main Survey revisions in their national version (using track changes) before submitting them for verification.

The verification of the New Mathematics coding guides was a focused verification only on revisions and concerned all CBA countries which had previously translated/adapted the Field Trial guides. Similar to the Main Survey verification of the cognitive units, countries could request changes either to correct residual errors or, in some cases, to modify the scoring instructions based on coder feedback or because the item showed differential functioning in the Field Trial, and a potential reason for this had been identified in the scoring instructions. If the National Centre did not request changes in the trend guides, these were not verified at all and the few revisions from Field Trial to Main Survey in Trend were left under National Centre responsibility.

The Main Survey verification procedure of coding guides was similar to that of cognitive items and followed the same workflow on the portal: countries could request justified changes to trend in the "Coding guide feedback form" in Excel format (CFF). The main difference compared to cognitive units was that all changes were implemented by the countries, while for cognitive units the countries did not have access to the files at any point, and verifiers made the changes in their New and Trend guides.

The translation memories from their final cognitive instruments were included in the national OmegaT packages, thus the quotations from the test items were identical with the instruments. The translation memories from their Field Trial coding guides were also included, and for the source segments that stayed identical as in the Field Trial, the translation was auto populated. The target segments for which the source segments changed in Main Survey were empty, while the translation from the Field Trial was available in the fuzzy matches pane. The country could update the Main Survey coding guides and correct the errata using the existing translation, as well as the consistency tools in the OmegaT. For the adapted versions, Chinese and Spanish Main Survey common reference versions were produced, and their translation memories from the Main Survey instruments and Field Trial coding were included in their national packages. These countries had to make sure that their adaptations were correctly reflected in the updated segments.

The completed forms and revised XLIFF and Word files were then submitted to Translation Referee for approval. Once the Referee had finished the review of the CFF, the files moved to verification. For the New Mathematics coding guide, the Referee review took place after verification. The verification and Referee review outcomes were documented in the same CFF. At verification, the DIF report was checked to make sure no undocumented changes were made.

When the National Centre did not request any changes to trend, a spot check was performed to their coding guides. If such changes were discovered the National Centre was asked either to provide a complete documentation, or to start over the preparation of the Main Survey guides (for example, if by mistake an outdated version was used as starting point).

For the countries that decided to use the master version as such either in ENG or FRA (e.g., Germany), the guides were not verified.

Errata management during Main Survey

Errata in Cognitive materials

Before the Main Survey preparation started, an all-in-one Field Trial to Main Survey Errata Document was released to the countries. This document included the errata released after the Field Trial and during the Main Survey review process for the Cognitive units. Countries/economies did not need to request the implementation of Main Survey errata. All of these errata were systematically checked and corrected by the verifiers, at verification step. At Post-Verification review step, the Countries/economies had to make sure that all released Main Survey errata have been addressed in a satisfactory way. If any Main Survey erratum was missed at verification, Countries/economies needed to indicate this in the CFF Coding Guides Follow-up Form, providing 1) the errata reference (from col. "Reference") and 2) the corrected version that the verifier should implement (whole segment). It was then addressed at Final Check.

The errata list included separate lists of errata identifying the errors in the English and the French source, as well as a separate tab with one erratum to be corrected in the source version in French in trend Reading item R549Q12: the wrong option was deleted in Source after selection for PISA 2018. This did not apply to National French versions for countries who participated to PISA 2018. The NCs were instructed to refer to that document to double-check if any of the errors listed in that file affected their national version if the reconciler had relied on the translation produced from French for a particular unit or section.

Errata in Questionnaires

The errata that were identified and approved for correction by the contractors before the Main Survey were documented in the Questionnaire Change Request Form in the Main Survey QAS, and the Countries/economies provided the corrected version in it. The verifiers then implemented the correction at verification step. The Countries/economies checked that the implementation was correct and documented residual issues, addressed by verifier at final check.

Suggestions for the future

The suggestions and lessons learnt in the PISA 2018 were taken on board and the process was significantly improved in PISA 2022. The major break-through in PISA 2022 was the use of OmegaT for translation, adaptation and verification of the PISA instruments. The PISA 2022 portal presented a clearer overview, a straightforward layout and yet a number of improved functionalities over the previous cycle. The coding guides for the new cognitive units were translated and verified in XLIFF format in OmegaT, benefitting from the translation memories from the verified cognitive units. The questionnaires were adapted in QAT, national master in XLIFF was exported from the QAT and translated in OmegaT. The Main Survey procedures for the cognitive and questionnaires got closer – a Questionnaire Change Request Form was used in the Main Survey verification.

At the conclusion of this process, we have the following specific recommendations in three areas.

Communication with countries and processes

In this cycle, communication with countries worked well. The trainings and webinars, the video tutorials, the User Guides, the questions, and answers section on the portal all contributed to clarify the different tasks to be performed at country level. In addition, at the end of each step, the NPM received an email with the instructions for the next step. On the other hand, not all national centres consulted and followed the instructions as expected. This could be due to various factors, such as (i) national centres not finding the instructions, (ii) national centre delegating the task to a person without forwarding the instructions (iii) user not understanding the instructions. The complexity of the PISA procedures and workflows may be rendered more understandable to the users if they are explained in pre-recorded webinar sessions that the

Countries/economies should watch before the live sessions during the face-to-face trainings and/or live webinar sessions. The trainings and the live sessions would then focus on Countries/economies' questions, issues, hands-on exercises, and particular difficulties.

File management

Although the PISA 2022 CBA Countries/economies could benefit from powerful translation memory management of the open-source CAT tool OmegaT in PISA 2022, version management issues were still a challenge in this cycle, i.e., national centre uploading an outdated version back to the workflow and pushing it forward, or national centre editing an outdated version and pushing it forward, losing the feedback provided in a previous step. A team OmegaT project may resolve this issue, where the online OmegaT package is automatically opened at each step of the workflow.

Errata management

Although in this cycle the errata management process was improved over PISA 2018, it was still observed that corrections were not implemented in the materials by the national centres. In the next cycle the errata management could also benefit from the use of OmegaT team project approach: at each source update, the target segments would appear untranslated, and the existing (outdated) translation from the translation memory would be shown in the fuzzy matches for reference. The user would then need to correct the translation so that it matches the updated source version.

Chapter 7 tables

Tables	Title
Table 7.1	Verified language versions of the PISA 2022 materials
Table 7.2	List of New Mathematics units in PISA 2022 Field Trial
Table 7.3	List of Field Trial New Mathematics not administered in the Main Survey
Table 7.4	List of Creative Thinking units in PISA 2022 Field Trial
Table 7.5	Translation procedures reported by national centres in the translation plan
Table 7.6	Example of Item Feedback Form

Table 7.1. Verified language versions of the PISA 2022 materials

PISA Participant	Language	Code	Last Cycle	Mode	PBA->CBA	Adpt	CT	FL	ICQ	TCQ	WBQ	UH	PAQ
Albania	Albanian	sqi-ALB	2018	CBA			Y**		Y	Y			
Argentina	Spanish	esp-ARG	2018	CBA	Y	Y			Y				
Australia	English	eng-AUS	2018	CBA		Y	Y		Y	Y			
Austria	German	deu-AUT	2018	CBA		Y			Y			Y	
Azerbaijan (Baku city only)	Azerbaijani	aze-QAZ	2018	CBA			Y			Y			
Azerbaijan (Baku city only)	Russian	rus-QAZ	2018	CBA		Y	Y			Y			
Belgium	French	fra-BEL	2018	CBA		Y	Y		Y			Y	
Belgium	Dutch	nld-BEL	2018	CBA		Y	Y		Y				Y
Bosnia and Herzegovina	Bosnian	bos-BIH	2018	CBA		Y	Y					Y	
Bosnia and Herzegovina	Croatian	hrv-BIH	2018	CBA		Y	Y					Y	
Bosnia and Herzegovina	Serbian	srp-BIH	2018	CBA		Y	Y					Y	
Brazil	Portuguese	por-BRA	2018	CBA			Y	Y	Y	Y	Y		Y
Brunei Darussalam	English	eng-BRN	2018	CBA		Y	Y		Y				
Bulgaria	Bulgarian	bul-BGR	2018	CBA			Y	Y	Y				
Cambodia	Khmer	khm-KHM	PISA-D	PBA									
Canada	English	eng-CAN	2018	CBA		Y	Y	Y					
Canada	French	fra-CAN	2018	CBA		Y	Y	Y					
Chile	Spanish	esp-CHL	2018	CBA		Y	Y	Y	Y	Y	Y	Y	Y
B-S-J-Z (China)	Chinese (simpl.)	zho-CHN	2018	CBA		Y	Y		Y	Y			
Colombia	Spanish	esp-COL	2018	CBA		Y	Y			Y			Y
Costa Rica	Spanish	esp-CRI	2018	CBA		Y	Y	Y	Y	Y	Y	Y	Y
Croatia	Croatian	hrv-HRV	2018	CBA			Y		Y				Y
Cyprus	Greek	ell-QCY	2018	CBA		Y	Y						
Cyprus	English	eng-QCY	2018	CBA		Y	Y						
Czech Rep.	Czech	ces-CZE	2018	CBA			Y	Y	Y			Y	
Denmark	Danish	dan-DNK	2018	CBA			Y	Y	Y			Y	
Dominican Republic	Spanish	esp-DOM	2018	CBA		Y	Y		Y	Y			Y
EI Salvador	Spanish	esp-SLV	NEW	CBA		Y	Y						

PISA Participant	Language	Code	Last Cycle	Mode	PBA->CBA	Adpt	CT	FL	ICQ	TCQ	WBQ	UH	PAQ
Estonia	Estonian	est-EST	2018	CBA			Y		Y				
Estonia	Russian	rus-EST	2018	CBA		Y	Y		Y				
Finland	Finnish	fin-FIN	2018	CBA			Y		Y				
France	French	fra-FRA	2018	CBA		Y	Y				Y		
Georgia	Georgian	geo-GEO	2018	CBA					Y	Y			Y
Germany	German	deu-DEU	2018	CBA			Y		Y	Y		Y	Y
Greece	Greek	ell-GRC	2018	CBA			Y		Y				
Guatemala	Spanish	esp-GTM	PISA-D	PBA		Y							
Hong Kong (China)	Chinese (trad.)	zho-HKG	2018	CBA		Y	Y		Y	Y	Y		Y
Hungary	Hungarian	hun-HUN	2018	CBA			Y	Y	Y		Y		
Iceland	Icelandic	isl-ISL	2018	CBA			Y		Y			Y	
India (Chandigarh)	English	eng-QIN	NEW	PBA		Y							
India	Hindi	hin-QIN	NEW	PBA									
Indonesia	Bahasa Indonesia	ind-IDN	2018	CBA			Y	Y					
Ireland	English	eng-IRL	2018	CBA		Y			Y		Y		Y
Israel	Arabic	ara-ISR	2018	CBA			Y		Y				
Israel	Hebrew	heb-ISR	2018	CBA			Y		Y				
Italy	Italian	ita-ITA	2018	CBA			Y	Y	Y				Y
Jamaica	English	eng-JAM	NEW	CBA		Y	Y						Y
Japan	Japanese	jpn-JPN	2018	CBA						Y			
Jordan	Arabic	ara-JOR	2018	CBA	Y		Y		Y				
Kazakhstan	Kazakh	kaz-KAZ	2018	CBA			Y		Y				
Kazakhstan	Russian	rus-KAZ	2018	CBA		Y	Y		Y				
Korea	Korean	kor-KOR	2018	CBA			Y		Y	Y	Y		Y
Kosovo	Albanian	sqi-KSV	2018	CBA		Y				Y			
Latvia	Latvian	lav-LVA	2018	CBA			Y		Y				Y
Latvia	Russian	rus-LVA	2018	CBA			Y	Y		Y			Y
Lebanon	English	eng-LBN	2018	CBA	Y	Y	Y		Y	Y			
Lebanon	French	fra-LBN	2018	CBA	Y	Y	Y		Y	Y			
Lithuania	Lithuanian	lit-LTU	2018	CBA			Y		Y				
Macao (China)	English	eng-MAC	2018	CBA		Y	Y		Y	Y	Y		Y
Macao (China)	Chinese (trad.)	zho-MAC	2018	CBA		Y	Y		Y	Y	Y		Y
Malaysia	Malaysian	msa-MYS	2018	CBA			Y	Y	Y	Y			
Malaysia	English	eng-MYS	2018	CBA			Y	Y	Y	Y	Y		
Malta	English	eng-MLT	2018	CBA			Y	Y		Y			
Malta	Maltese	mlt-MLT	2018	CBA			Y			Y			
Mexico	Spanish	esp-MEX	2018	CBA		Y	Y			Y	Y		
Moldova	Romanian	ron-MDA	2018	CBA	Y	Y	Y						
Moldova	Russian	rus-MDA	2018	CBA	Y	Y	Y						
Mongolia	Mongolian	mon-MNG	NEW	CBA			Y						
Montenegro	Montenegrin	mne-MNE	2018	CBA		Y							

PISA Participant	Language	Code	Last Cycle	Mode	PBA->CBA	Adpt	CT	FL	ICQ	TCQ	WBQ	UH	PAQ
Morocco	Arabic	ara-MAR	2018	CBA			Y		Y	Y			
Morocco	French	fra-MAR	2018	CBA		Y	Y						
Netherlands	Dutch	nld-NLD	2018	CBA			Y	Y			Y	Y	
New Zealand	English	eng-NZL	2018	CBA		Y	Y		Y		Y		
North Macedonia	Albanian	sqi-MKD	2018	CBA	Y	Y	Y						
North Macedonia	Macedonian	mkd-MKD	2018	CBA	Y		Y						
Norway	Bokmål	nob-NOR	2018	CBA				Y				Y	
Norway	Nynorsk	nno-NOR	2018	CBA		Y		Y				Y	
Panama	Spanish	esp-PAN	2018	CBA		Y	Y		Y	Y	Y		Y
Paraguay	Spanish	esp-PRY	PISA-D	PBA		Y							
Peru	Spanish	esp-PER	2018	CBA		Y	Y	Y		Y			
Philippines	English	eng-PHL	2018	CBA		Y	Y						
Poland	Polish	pol-POL	2018	CBA			Y	Y	Y				
Portugal	Portuguese	por-PRT	2018	CBA			Y	Y		Y			Y
Qatar	Arabic	ara-QAT	2018	CBA			Y						
Qatar	English	eng-QAT	2018	CBA		Y	Y						
Romania	Romanian	ron-ROU	2018	CBA	Y		Y		Y				
Saudi Arabia	Arabic	sau-ARA	2018	CBA	Y		Y						Y
Serbia	Serb (Ekavian)	srp-SRB	2018	CBA			Y	Y					
Singapore	English	eng-SGP	2018	CBA		Y	Y		Y				
Slovak Rep.	Slovak	slo-SVK	2018	CBA			Y		Y				Y
Slovenia	Slovenian	slv-SVN	2018	CBA			Y		Y		Y	Y	
Spain	Basque	eus-ESP	2018	CBA			Y	Y	Y			Y	
Spain	Galician	glg-ESP	2018	CBA			Y	Y	Y			Y	
Spain	Castilian	esp-ESP	2018	CBA		Y	Y	Y	Y			Y	
Spain	Catalan	cat-ESP	2018	CBA			Y	Y	Y			Y	
Sweden	Swedish	swe-SWE	2018	CBA						Y			
Switzerland	French	fra-CHE	2018	CBA		Y			Y				
Switzerland	German	deu-CHE	2018	CBA		Y			Y				
Chinese Taipei	Chinese (trad.)	zho-TAP	2018	CBA		Y	Y		Y				
Thailand	Thai	tha-THA	2018	CBA			Y		Y				
Turkey	Turkish	tur-TUR	2018	CBA			Y		Y				
Ukraine	Ukrainian	ukr-UKR	2018	CBA	Y		Y		Y				Y
UAE	Arabic	ara-ARE	2018	CBA			Y	Y	Y		Y	Y	
UAE	English	eng-ARE	2018	CBA		Y	Y	Y		Y	Y		
United Kingdom (excl. Scotland)	English	eng-QUK	2018	CBA		Y			Y				
United Kingdom (Scotland)	English	eng-QSC	2018	CBA		Y	Y		Y				
United States	English	eng-USA	2018	CBA		Y		Y	Y	Y			Y

PISA Participant	Language	Code	Last Cycle	Mode	PBA->CBA	Adpt	CT	FL	ICQ	TCQ	WBQ	UH	PAQ
Uruguay	Spanish	esp-URY	2018	CBA		Y	Y		Y				
Uzbekistan	Uzbek	uzb-UZB	NEW	CBA			Y						
Vietnam	Vietnamese	vie-VNM	2018	PBA									

Note:

* This list reflects countries and economies that submitted instruments for verification. For actual participation status, please refer to Table 1.1 in this report.

** "Y" stands for "Yes" in this table.

Table 7.2. List of New Mathematics units in PISA 2022 Field Trial

Batch	Unit identifier	Unit
Batch 1	MA101	Building Blocks
	MA102	Buying a Wardrobe
	MA103	Calculation Program
	MA104	Car Purchase
	MA105	Clean Energy
	MA106	DVD Sales
	MA107	Field OF Vision
	MA108	Fountains
	MA109	Headache Medicine
	MA112	Metabolism
	MA125	Painting A Room
	MA128	Salinity OF Water
	MA153	Gears
	MA159	Spinners
	MA160	University Student Employment
	MA161	Forested Areas
	MA162	Urban Population
Batch 2	MA110	Headphone Order
	MA111	Health App
	MA113	Heart Rate
	MA114	Honey
	MA115	Iceberg
	MA116	International School
	MA117	Mixing Paint
	MA118	Moving Truck
	MA119	Music Survey
	MA120	Number Cubes
	MA121	Mobile Phone Reviews
	MA122	Pool Cover
	MA123	Solar System
	MA124	Zedland Topography
	MA126	Robot

Batch	Unite identifier	Unit
	MA127	Predicting Height
	MA129	Shelving Unit
	MA130	Sleep and Reaction Time
	MA131	Travelling by Train
	MA132	Water Temperature
Batch 3	MA133	Arranging Tables
	MA134	Car and Bicycle Ownership
	MA135	Electric Bicycle
	MA136	Movie Rewards
	MA137	Football Tournament
	MA138	Shoe Sizes
	MA139	Tablet Cover
	MA140	Walk to School
	MA141	Water Bill
	MA142	Water Reservoir
	MA143	Wild Bird Food
	MA144	Yogurt
	MA145	Shadows
	MA146	Fuel
	MA147	Aeroplane Tickets
	MA148	Chance of Rain
	MA149	Floor Area
	MA150	Triangular Pattern
	MA151	Moving Out
	MA152	The Better Deal
	MA154	Company Logo
	MA156	Points
	MA157	Tyres
	MA158	Eye Colour
Batch 6A	M905	Tennis Balls
	M919	Fan Merchandise
	M943	Arches
	M953	Flu Test
	M954	Medicine Doses
Batch 6B	M936	Seats in a Theatre
	M939	Racing
	M948	Part Time Work
	M961	Chocolate
	M967	Wooden Train Set

Table 7.3. List of Field Trial New Mathematics not administered in the Main Survey

New Mathematics	Dropped item/unit in MS
MA101 Q03	Dropped item
MA103 Q03	Dropped item
MA104	Dropped unit
MA106	Dropped unit
MA114 Q02	Dropped item
MA117 Q05	Dropped item
MA118	Dropped unit
MA122	Dropped unit
MA123 Q03	Dropped item
MA126 Q01	Dropped item
MA136 Q01	Dropped item
MA137 Q02	Dropped item
MA144 Q02	Dropped item
MA156 Q02	Dropped item
MA159	Dropped unit

Table 7.4. List of Creative Thinking units in PISA 2022 Field Trial

Unit Identifier	Unit
T200	Science Fair Poster
T240	Space Comic
T300	Illustration Titles
T350	Book Covers
T360	Moving Backward
T370	2983
T400	Save the Bees
T420	Clean Oceans
T450	Music Festival
T500	Wheelchair Accessible Library
T520	Painting Class
T540	Infographics
T550	Experiment Kit
T560	The Ball
T570	Robot Story
T610	Food Waste
T620	Paper Products
T630	Carpooling
T680	Rubber Ducks Game
T690	Save the River
T700	The Exhibit

Table 7.5. Translation procedures reported by national centres in the translation plan

Type	Cognitive Items	Questionnaires
Double translation from English and French source versions	17	18
Double translation from English source version with cross-checks against the FRA source version	8	8
Double translation from English source version only	30	39
Adaptation from one of the source versions	25	25
Adaptation from a borrowed verified version or from a common base version	29	23
Double translation from English source version with cross-checks against the Spanish common reference version	3	2

Table 7.6. Example of Item Feedback Form

Refer to Excel file <XXX> to view this table on line.

Notes

1. A translation memory is a database that stores sentences, paragraphs or segments of text that have been translated before.
2. Following Note 4.1 to the PISA 2022 Technical Standards.

Annex 7.A. Verifier intervention categories

Annex Table 7.A.1. Verifier intervention categories

Category	Description
OK	No intervention is needed. The verifier has checked and confirms that the text element or segment is equivalent to source, linguistically correct, and – if applicable – that it conforms to an explicit translation/adaptation guideline. This category may also be used to report an appropriate but undocumented adaptation.
Added information	An information is present in the target version but not in the source version, e.g., an explanation between brackets of a preceding word.
Missing information	An information is present in the source version but omitted in the target version.
Matches and patterns	A literal match (repetition of the same word or phrase) or a synonymous match (use of a synonym or paraphrase) in the source version is not reflected in the target version. Most important: literal or synonymous matches between stimulus and item and between a question stem and response categories. A pattern in multiple choice items is not reflected in the target version (e.g., all but one option starts with the same word, proportional length of responses options.)
Inconsistency	A recurring element across units (e.g., an instruction or prompt) is inconsistently translated, and this appears to be unintentional.
Adaptation issue	An adaptation is an intentional deviation from the source version made for cultural reasons or to conform to local usage. An adaptation issue occurs when an adaptation would be needed but was not made, or when an inappropriate or unnecessary adaptation was made.
Register / Wording issue	<i>Register:</i> difference in level of terminology (scientific term >< familiar term) or level of language (formal >< casual, standard >< idiomatic) in target versus source. <i>Wording:</i> inappropriate or less than optimal choice of vocabulary or wording in target to fluently convey the same information as in the source. This category is used typically for vague or inaccurate or not quite fluent translations.
Grammar / Syntax issue	<i>Grammar:</i> grammar mistake that could affect comprehension or equivalence, e.g., wrong subject-verb agreement, wrong case (inflected languages), wrong verb form. <i>Syntax:</i> syntax-related deviation from the source, e.g., a long (source) sentence is split into two (target) sentences or two (source) sentences are merged into a single (target) one; or another syntactic problem due e.g., to overly literal translation of the source.
Mistranslation	A wrong translation, which seriously alters the meaning. A <u>mistranslation should always be reported with a back-translation</u> . Note: a vague or inaccurate translation should rather be classified as a Register/Wording issue (or sometimes a Grammar/Syntax issue). This category covers cases where the source has been misunderstood, but also copy/paste errors that unintentionally result in a wrong text element or segment.
Guideline not followed	An explicit translation/adaptation guideline for a given text element or segment was overlooked or was not addressed in a satisfactory way.
Left in source language	A text element or segment that should have been translated was left in source language.
Minor linguistic issue	Typo or other linguistic defect (spelling, grammar, capitalization, punctuation, etc.) that does not significantly affect comprehension or equivalence. Correcting such errors is usually not controversial and can be made in track changes without documenting them.
Erratum/Update missed	An erratum or update notice has been overlooked.
Layout / Format issue	A deviation or defect in layout or formatting: disposition of text and graphics, item labels, question numbering, styles (boldface, <u>underlining</u> , <i>italics</i> , UPPERCASE), legibility of captions, tables, number formatting (decimal separators, “five” versus “5”), etc. In CB materials, this includes truncated words in the preview, undesired scrolling, etc.

Annex 7.B. Translatability Assessment categories

Annex Table 7.B.1. Translatability Assessment categories

Category	Description
Straightforward	No potential translation or adaptation problems identified during the advance translation of this segment into languages from at least two language groups.
Known difficulty, known workarounds	A translation/adaptation difficulty has been recognised in this segment and has been encountered in the past. Satisfactory solutions to this issue have been successfully implemented.
Potential issues	The current wording or content of this segment is likely to give rise to translation or adaptation problems in some languages, to the extent that functional equivalence may be difficult to achieve.
Potentially ambiguous	The current wording or content of this segment could be interpreted in more than one way and it is desirable to disambiguate the source version of this segment before submitting it for translation/adaptation.
Unnecessarily complex	The current wording or syntax of this segment is somewhat contorted, for example due to use of several clauses, questions embedded in questions or unnecessary use of passive voice. The source version can be simplified without loss of meaning.
Requires review	The current source version of this segment is not suitable for translation/adaptation and needs to be edited before submitting for translation/adaptation.
Potential cultural issue	The semantic content of this segment may be difficult to adapt in a particular cultural or language group.
Double-barrelled	A question touches upon more than one issue, yet allows only for one answer. Many double-barrelled questions can be detected by the existence of the grammatical conjunction "and" in them.
Agreement issue	There is either an agreement issue within the segment (e.g., subject-verb agreement, or sequence of tenses, or a pronoun-antecedent agreement) or an agreement issue between two segments (e.g., no grammatical match between a question and response options).
Consistency	In this segment, a different term, expression or form of address has been used versus other occurrences of similar content; and this inconsistency seems to be unintentional.
Redundancy	This segment contains a tautology or unnecessary repetition. Removing it would not alter the meaning of the segment.
Possible addition	The current wording or syntax of this segment is elliptic or unclear, and its implicit meaning is likely to get lost in translation. This could be solved by adding a word or a piece of information.
Logical problem	This segment contains a logical problem or there is a logical problem between this segment and another segment, and this issue seems to be unintentional.

Annex 7.C. Translation Plan

Refer to Excel file Appendix_3_CY8_GEN_TranslationPlan_2021.xlsx

Annex 7.D. Verification outcomes in New Mathematics units per language version

Refer to Excel file <Appendix_4_PISA2022_New-MAT_per_language.xlsx> for detailed statistics

Annex 7.E. Verification outcomes in New Mathematics units per cognitive unit

Refer to Excel file <[Appendix_5_PISA2022_New-MAT_per_unit.xlsx](#)> for detailed statistics

Annex 7.F. Verification outcomes in Creative Thinking units per language version

Refer to Excel file <Appendix_6_PISA2022_CreativeThinking_per_language.xlsx> for detailed statistics

Annex 7.G. Verification outcomes in Creative Thinking units per cognitive unit

Refer to Excel file <Appendix_7_PISA2022_CreativeThinking_per_unit.xlsx> for detailed statistics

This work is published under the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of the Member countries of the OECD.

Note by the Republic of Türkiye

The information in this document with reference to "Cyprus" relates to the southern part of the Island. There is no single authority representing both Turkish and Greek Cypriot people on the Island. Türkiye recognises the Turkish Republic of Northern Cyprus (TRNC). Until a lasting and equitable solution is found within the context of the United Nations, Türkiye shall preserve its position concerning the "Cyprus issue".

Note by all the European Union Member States of the OECD and the European Union

The Republic of Cyprus is recognised by all members of the United Nations with the exception of Türkiye. The information in this document relates to the area under the effective control of the Government of the Republic of Cyprus.

The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.

The use of this work, whether digital or print, is governed by the Terms and Conditions to be found at:
<https://www.oecd.org/termsandconditions>