

ANNEX D – TECHNICAL STANDARDS FOR PISA FOR DEVELOPMENT

INTRODUCTION

The purpose of this annex is to list the set of standards upon which the PISA for Development (PISA-D) data collection activities were based. In following the procedures specified in the standards, the partners involved in the data collection activities contributed to creating an international dataset of a quality that allows for valid cross-national inferences to be made. In order to ensure comparability between PISA and PISA-D, this document relies on the standards for PISA 2018 as the starting point and identifies those key aspects within each standard that were modified for PISA-D in the following ways:

- Text that does not apply to PISA-D is noted in a lighter (grey) font.
- Text has been added to reflect the PISA-D context as appropriate.

The standards for data collection and submission were developed with three major, and inter-related, goals in mind: consistency, precision, and generalisability of the data. Furthermore, the standards serve to ensure a timely progression of the project in general.

- **Consistency:** Data should be collected in an equivalent fashion in all countries, using equivalent test materials that were translated and/or adapted as appropriate. Comparable samples of each country’s student population should perform under test conditions that are as similar as possible. Given consistent data collection (and sufficiently high response rates), test results are likely to be comparable across regions and countries. The test results in different countries will reflect differences in the performance of the students measured, and will not be caused by factors which are unrelated to performance.
- **Precision:** Data collection and submission practices should leave as little room as possible for spurious variation or error. This holds for both systematic and random error sources, for example, when the testing environment differs from one group of students to another, or when data entry procedures leave room for interpretation. An increase in precision relates directly to the quality of results one can expect: The more precise the data, the more powerful the (statistical) analyses, and the more trustworthy the results to be obtained.
- **Generalisability:** Data are collected from specific individuals, in a specific situation, and at a certain point in time. Individuals to be tested should be selected, and test materials and tasks, and so on, be developed in a way that will ensure that the conclusions reached from a given set of data do not simply reflect the setting in which the data were collected but hold for a variety of settings and are valid for the target population at large. Thus, collecting data from a representative sample of the population, for example, will be essential, but not sufficient, for the results to accurately reflect the level of literacy of 15-year-old students in a country.

- **Timeliness:** Consistency, precision, and generalisability of the data can be obtained in a variety of ways. However, the tight timelines and budgets in PISA-D preclude the option of developing and monitoring local solutions to be harmonised at a later stage in the project. Therefore, the standards specify guidelines for data collection and data submission.

These standards strive to establish a collective agreement of mutual accountability among countries, and of the international contractors towards the countries. This document details each standard and the quality assurance and quality management plan to demonstrate that the standard has been met. While the terms “quality assurance” and “quality control” are sometimes used interchangeably, they relate to different aspects of quality. Quality assurance is most often associated with the processes and procedures that are put in place to make sure the survey is likely to meet its intended goals. Quality control, on the other hand, relates to the set of judgements that are made with regard to the suitability of the quality assurance procedures and the suitability of the survey results in terms of their intended uses or applications.

Where standards have been fully met and data quality of the final database judged as appropriate, the international contractors have recommended to the OECD Secretariat that the data be included in the PISA-D database. Where standards have not been fully met or data quality has been questioned, an adjudication process has determined the extent to which the quality and international comparability of the data have been affected or whether additional analysis or evidence were necessary. The result of data adjudication determined whether the data was recommended for inclusion in the PISA-D dataset.

In principle, each dataset should be evaluated against all standards jointly. Also, it is possible that countries’ proposed plans for implementation were not, for various and often unforeseen circumstances, actually implemented (e.g. national teacher strike affecting not only response rates, but also testing conditions; unforeseen National Centre budget cuts that affect print and data management quality). Therefore, the final evaluation of standards was made with respect to the data as submitted since this is the definitive indication of what appears in the released international dataset.

When any issues with attaining standards or data quality were identified, the International Survey Director initiated communication with the National Centre as soon as possible to give advice on resolving problems.

The PISA-D standards serve as benchmarks of best practice. As such, the standards are designed to assist National Centres and international contractors by explicitly indicating the expectations of data quality and study implementation endorsed by the PISA Governing Board, and by clarifying the timelines of the activities involved. The standards formulate levels of attainment, while timelines and feedback schedules of both the participating countries and the international contractors are defined in the PISA-D operations manuals.

As specified in the contracts for the implementation of the OECD Programme for International

Student Assessment for Development, the international contractors take responsibility for developing and implementing procedures for assuring data quality. Therefore, the international contractors mediate and monitor the countries' activities specified in this document.

Where the technical standards stipulate that variations from the standards require agreement between participating countries and the international contractors, National Project Managers were asked to initiate the process of negotiation and to undertake everything possible to facilitate an agreement. Where agreement between National Project Managers and the international contractors could not be reached, the OECD adjudicated and resolved the issues. The OECD also adjudicated any issues resulting from non-compliance with the technical standards that could not be resolved between participating countries and the contractors.

There are three types of standards in this document, each with a specific purpose:

- **Data Standards** refer to aspects of study implementation that directly concern the data quality and its control. These standards have been endorsed by the Technical Advisory Group and wherever proportions or quantities are specified (for example, response rates), have been reached through examination of research undertaken or have been reviewed by members of the Technical Advisory Group with the aim of minimising the effect of any potential bias in the data.
- **Management Standards** are in place to ensure that all PISA-D operational objectives are met in a timely and coordinated manner.
- **National Involvement Standards** reflect the expectation set out in the PISA-D Terms of Reference that the content of the PISA-D tests is established in consultation with national representatives with international content expertise. In particular, these standards help to ensure that the internationally developed instruments are widely examined for cross-national, cross-cultural, and cross-linguistic validity and that the interests and involvement of national stakeholders are considered throughout the study.

FORMAT OF THIS DOCUMENT

The standards are grouped into sections that relate to specific tasks in the PISA-D data collection process. For every section, a rationale is given explaining why standard setting is necessary. The standards in each section consist of three distinct elements. First, there are the **Standards** themselves that are numbered and are shown in shaded boxes. Second, there are **Notes** that provide additional information on the standards directly. The notes are listed after the standards in each section. Third, there are the **Quality Control** measures that are used to assess if a standard has been met or not. These are listed at the end of each section. In addition, the standards contain words that have a defined meaning in the context of the standards. These words are shown in italics throughout the document and are clarified in the **Definitions** section at the end of the document, where the terms are listed alphabetically.

SCOPE

The standards in this document apply to data from adjudicated entities that include PISA-D participants.

DATA STANDARDS

1. Target population and sampling

Rationale: Meeting the standards specified in this section will ensure that in all countries, the students tested come from the same target population in every country, and are in a nearly equivalent age range. Therefore, the results obtained were not confounded by potential age effects. Furthermore, to be able to draw conclusions that are valid for the entire population of 15-year-old students, a representative sample was selected for participation in the test. The size of this representative sample should not be too small, in order to achieve a certain precision of measurement in all countries. For this reason, minimum numbers of participating students and schools were specified.

The procedures for drawing the samples used in the study are crucial to data quality. The goal of the project is to collect data that are representative for the population at large in such a way that the results are comparable, reliable, and valid. To reach these goals the sampling procedures must follow established scientific principles for drawing samples from finite populations.

Standard 1.1 The *PISA Desired Target Population* is *agreed upon* through negotiation between the National Project Manager and the international contractors within the constraints imposed by the definition of the *PISA Target Population*.

Standard 1.2 Unless *otherwise agreed upon* only *PISA-D eligible students* participate in the test.

Standard 1.3 Unless *otherwise agreed upon*, the *testing period*:

- is no longer than eight consecutive weeks in duration for computer-based testing participants,
- is no longer than six consecutive weeks in duration for paper-based testing participants; and
- does not coincide with the first six weeks of the academic year; and
- begins exactly three years from the beginning of the *testing period* in the previous PISA cycle

Standard 1.4 Schools are sampled using *agreed upon*, established and professionally recognised principles of scientific sampling.

Standard 1.5 Student lists should not be collected more than 8 weeks prior to the start of data collection, unless *otherwise agreed upon*.

Standard 1.6 Students are sampled using *agreed-upon*, established and professionally recognised principles of scientific sampling and in a way that represents the full population of *PISA-Eligible students*.

Standard 1.7 The *PISA Defined Target Population* covers 95% or more of the *PISA Desired Target Population*. That is, *school-level exclusions* and *within-school exclusions* combined do not exceed 5%.

Standard 1.8 The student sample size for the **computer-based mode** is a minimum of 6,300 assessed students, and 2,100 for *additional adjudicated entities*, or the entire *PISA Defined Target Population* where the *PISA Defined Target Population* is below 6,300 and 2,100 respectively. The student sample size of assessed students for the **paper-based mode** is a minimum of 5,250.

Standard 1.9 The school sample size needs to result in a minimum of 150 participating schools, and 50 participating schools for *additional adjudicated entities*, or all schools that have students in the *PISA Defined Target Population* where the number of schools with students in the *PISA Defined Target Population* is below 150 and 50, respectively. Countries not having at least 150 schools, but which have more students than the required minimum student sample size, can be permitted, *if agreed upon*, to take a smaller sample of schools while still ensuring enough sampled PISA-D students overall.

Standard 1.10 The final weighted school response rate is at least 85% of sampled eligible and non-excluded schools. If a response rate is below 85%, then an acceptable response rate can still be achieved through *agreed-upon* use of replacement schools. To calculate the weighted school response rate, each eligible school is weighted by the product of the school base weight (the reciprocal of the school's selection probability) and the enrolment size of eligible students (as indicated on the sampling frame). The denominator of the weighted response rate is the sum of this quantity across all eligible sample schools. The numerator is the sum over all participating schools. Thus this rate indicates the proportion of the student population that is directly represented by the responding schools.

Standard 1.11 The final weighted student response rate is at least 80% of all sampled students across responding schools.

Standard 1.12 The final weighted sampling unit response rate for any optional cognitive assessment is at least 80% of all sampled students across responding schools.

Standard 1.13 Unless *otherwise agreed upon*, the international contractors will draw the school sample for the Main Survey.

Standard 1.14 Unless *otherwise agreed upon*, the National Centre will use MyMWestra to draw the student sample, using the list of eligible students provided for each school.

Note 1.1 Standards 1.1 through 1.14 apply to the Main Survey but not the Field Trial.

Note 1.2 Data from schools where the (unweighted) student response rate is greater than 25% will be included in the PISA-D dataset.

Note 1.3 For the purpose of calculating school response rates, a participating school is defined as a sampled school in which more than 50% of sampled eligible, non-excluded students respond.

Note 1.4 Acceptable response rates obtained through the use of replacement schools are described in detail in the School Sampling Preparation Manual.

Note 1.5 Guidelines for acceptable exclusions that do not affect standard adherence are as follows:

- School-level exclusions that are exclusions due to geographical inaccessibility, extremely small school size, administration of PISA-D would be not feasible within the school, and other agreed-upon reasons and whose students total to less than 0.5% of the PISA Desired Target Population,
- School-level exclusions that are due to a school containing only students that would be within-school exclusions and that total to less than 2.0% of the PISA Desired Target Population, and
- Within-school exclusions that total to less than 2.5% of the PISA Desired Target Population – these exclusions could include, for example, students not able to do the test because of a functional disability.

Note 1.6 Principles of scientific sampling include, but are not limited to:

- The identification of appropriate stratification variables to reduce sampling variance and facilitate the computation of non-response adjustments.
- The incorporation of an agreed target cluster size of PISA-Eligible students: The minimum acceptable target cluster size is 25 students. In determining the target cluster size for a given country, or stratum within a country, it is necessary to ensure that the minimum sample size requirements for both schools and students will be met.

Note 1.7 Any exceptional costs associated with verifying a school sample taken by the National Centre, or a student sample selected other than by using MyMWestra, will be borne by the National Centre.

Note 1.8 Agreement with the international contractor on alternative methods of drawing samples will be subject to the principle that the sampling methods used are scientifically valid and consistent with PISA-D's documented sampling methods. Where a PISA-D participating country chooses to draw the school sample, the National Centre provides the international contractor with the data and documentation required for it to verify the correctness of the sampling procedures applied. Where a PISA-D participating country chooses not to use MyMWestra to draw the student sample, the National Centre provides the international contractor with the data and documentation required for it to verify the correctness of the sampling procedures applied.

Quality control

- Sampling procedures as specified in the PISA-D operations manuals
- School sample drawn by the international contractors (or if drawn by the National Centre, then verified by the international contractors)
- Student sample drawn through MyMWestra (or if drawn by other means, then verified by the international contractors)
- Sampling forms submitted to the international contractors
- Main Survey Review Form

2. Language of testing

Rationale: Using the language of instruction ensures analogous testing conditions for all students within a country, thereby strengthening the consistency of the data. It is assumed that the students tested have reached a level of understanding in the language of instruction that is sufficient to be able to work on the PISA-D test without encountering linguistic problems (see also the criteria for excluding students from the potential assessment due to insufficient experience in the language of assessment: *within-school exclusions*). Thus, the level of literacy in reading, mathematics, and science can be assessed without interference due to a critical variation in language proficiency.

Standard 2.1 The PISA test is administered to a student in a language of instruction provided by the sampled school to that sampled student.

If the language of instruction in the major domain is not well defined across the set of sampled students then, if *agreed upon*, a choice of language can be provided, with the decision being made at the student, school, or National Centre level. Agreement with the international contractor will be subject to the principle that the language options provided should be languages that are common in the community and are common languages of instruction in schools in that *adjudicated entity*.

If the language of instruction differs across domains then, if *agreed upon*, students may be tested using assessment instruments in more than one language on the condition that the test language of each domain matches the language of instruction for that domain.

Information obtained from the Field Trial will be used to gauge the suitability of using assessment instruments with more than one language in the Main Survey.

In all cases the choice of test language(s) in the assessment instruments is made prior to the administration of the test.

3. Field Trial participation

Rationale: The Field Trial gives countries the opportunity to try out the logistics of their test procedures and allows the international contractors to make detailed analyses of the items so that only suitable ones are included in the Main Survey.

Standard 3.1 *PISA-D participants* participating in the PISA-D Main Survey will have successfully implemented the Field Trial. Unless otherwise *agreed upon*:

- A Field Trial should occur in an assessment language if that language group represents more than 5% of the target population.
- For the largest language group among the target population, the Field Trial student sample should be a minimum of 200 students per item.
- For all other assessment languages that apply to at least 5% of the target population, the Field Trial student sample should be a minimum of 100 students per item.
- For *additional adjudicated entities*, where the assessment language applies to at least 5% of the target population in the entity, the Field Trial *student sample should be a minimum of 100 students per item*.

Note 3.1 The PISA Technical Standards for the Main Survey generally apply to the Field Trial, except for the Target Population standard, the Sampling standard, and the Quality Monitoring standard. For the Field Trial, a sampling plan needs to be agreed upon.

Note 3.2 The sample size for the Field Trial will be a function of the test design and will be set to achieve the standard of 200 student responses per item.

Note 3.3 Consideration will be given to reducing the required number of students per item in the Field Trial where there are fewer than 200 students in total expected to be assessed in that language in the Main Survey.

4. Adaptation of tests, questionnaires, and school-level materials

Rationale: In order to be able to assess how the performance in a country has evolved from one

PISA cycle to the other, the same instruments have to be used in all assessments. If instruments differ, then it is unclear whether changes in performance reflect changes in literacy, or whether they just mirror the variation in the test items. The same holds true for the assessment instruments that are used within a PISA-D cycle: To validly compare performance across countries, all assessment instruments and other survey materials have to be as equivalent as possible. In fact, it is of utmost importance to provide equivalent information to the students in all countries that take part in the study. Therefore, not only the assessment instruments, but also the instructions given to the students and the procedures of data collection, have to be equivalent. To achieve this goal, other individuals who play a key role in the data-collection process, that is, the test administrators, school coordinators, and school associates, received equivalent information and training in all participating countries.

Standard 4.1 The majority of test items used for linking are administered unchanged from their previous administration in PISA. The computer-based versions include instructions as to the appropriate response mode for each item and may require some minor revision as noted in 4.2 below.

Standard 4.2 All assessment instruments are psychometrically equivalent to the *source versions*. *Agreed-upon* adaptations to the local context are made if needed.

Standard 4.3 National versions of questionnaire items used in previous cycles will be administered unchanged from their previous administration, unless amendments have been made to source versions.

Standard 4.4 The questionnaire instruments are equivalent to the *source versions*. *Agreed-upon* adaptations to the local context are made if needed.

Standard 4.5 *School-level materials* are equivalent to the *source versions*. *Agreed-upon* adaptations to the local context are made if needed.

Note 4.1 The quality control requirements for this standard apply to instruments that are in an assessment language used as a language of instruction for more than 10% of the target population.

Quality control

- Agreed-upon adaptation to school-level materials using method specified by the international contractors
- Questionnaire Adaptation Spreadsheet (QAS)
- Test Adaptation Spreadsheets (TAS, for paper and computer instruments) and Booklet Adaptation Spreadsheets (BAS, for paper-based booklets only), in which adaptations to assessment units, orientation and help files, common booklet parts, and coding guides were documented. For languages that are the languages of instruction for 10% or more of the target population, adaptations were checked for compliance with the PISA-D Translation and Adaptation Guidelines by international verifiers and the verifiers' recommendations were vetted by the translation referee.

- For languages that are the languages of instruction for 10% or more of the target population: Verifier Reports (statistics generated by the TAS, in combination with a short qualitative report)
- Field Trial and Main Survey Review Forms
- Item and scale statistics generated by the international contractors (assessment materials and questionnaires)

5. Translation of assessment instruments, questionnaires, and school-level materials

Rationale: To be able to compare the performance of students across countries, and of students with different instruction languages within a country, the linguistic equivalence of all materials is central. While Standards 4.1 to 4.4 serve to ensure that equivalent information is given to the students in all countries involved, in general, the following Standards 5.1 and 5.2 emphasise the importance of language. Again the goal is to ensure that literacy is assessed, and not variations of information caused by differences in the translation of materials.

Standard 5.1 The following documents are translated into the assessment language in order to be linguistically equivalent to the international *source versions*.

- All administered assessment instruments
- All administered questionnaires
- The Test Administrator script from the Test Administrator (or School Associate) Manual
- The coding guides

Standard 5.2 Unless otherwise *agreed upon*, school-level materials are translated/adapted into the assessment language to make them functionally equivalent to the international *source versions*.

Note 5.1 The quality assurance requirements for this standard apply to instruments that are in a language that is administered to more than 10% of the target population.

Quality control

- Agreed-upon Translation Plan developed in accordance with the specifications in the PISA-D operations manuals that require double translation by independent translators followed by reconciliation
- Agreed-upon Questionnaire Adaptation Spreadsheet (QAS)
- Test Adaptation Spreadsheets (TAS) and Booklet Adaptation Spreadsheet (BAS, for paper-based booklets only), in which adaptations to assessment units, orientation and help files, common booklet parts, and coding guides were documented. Adaptations were checked for compliance with the PISA-D Translation and Adaptation Guidelines by international verifiers, and the verifiers' recommendations were vetted by the translation referee.
- Verifier Reports (statistics generated by the TAS, in combination with a short qualitative report)

- Submitted final materials as used in the study
- Field Trial and Main Survey Review Forms
- Item and scale statistics generated by the international contractors (assessment materials and questionnaires)

6. Testing of national software versions

Rationale: Countries must thoroughly test and validate the national software releases that are used to deliver the PISA computer-based instruments in schools, as well as the online questionnaires that are delivered via the Internet.

Standard 6.1 The international contractors must test all national software versions prior to their release to ensure that they were assembled correctly and have no technical problems.

Standard 6.2 Once released, countries must test the national software versions following testing plans to ensure the correct implementation of national adaptations and extensions, display of national languages, and proper functioning on computers typically found in schools in each country.

Note 6.1 Errors found during testing should be promptly communicated to the international contractors using agreed-upon problem reporting procedures. These procedures require that testing results are shared with the international contractors in order to monitor the quality of the instruments.

Quality assurance

- Detailed testing plans
- Review of testing results

7. Technical support

Rationale: Countries participating in the computer-based delivery mode will be primarily responsible for resolving PISA-related operational issues in their countries, including hardware issues and provision of technical support to schools and test administrators.

Standard 7.1 Each country should have a designated PISA helpdesk with contact information provided to each of its *test administrators* and school coordinators.

Standard 7.2 The country helpdesk staff must:

- be familiar with the PISA computer system requirements applications and training materials,
- be familiar with all national software standards and procedures; and
- attend the *test administrator* training sessions to become familiar with the computer-based assessments and appreciate the challenges faced by schools and *test administrators*.

Quality assurance

- National Centre Quality Monitoring
- Field Trial and Main Survey Review Forms

8. Test administration

Rationale: Certain variations in the testing procedure are particularly likely to affect test performance. Among them are session timing, the administration of test materials and support material like blank papers and calculators, the instructions given prior to testing, the rules for excluding students from the assessment, etc. A full list of relevant test conditions is given in the school-level materials. To ensure that the data are collected consistently, and in a comparable fashion, for all participants, it is therefore very important to keep the chain of action in the data collection process as constant as possible.

Furthermore, the goal of the assessment is to arrive at results which cover a wide range of areas. Given the time constraints, any one student is presented only with a certain portion of the test items. Moreover, to preclude sources of random error unforeseen by the test administrators and the test designers, the students taking part in the survey have to be selected *a-priori*, in a statistically random fashion. Only then do the students participating in the study mirror the population of 15-year-old students in the country. The statistical analysis takes this sampling design into account, thereby arriving at results that are representative for the population at large. For these reasons, it is of utmost importance to assign the proper test booklets or computer equivalent to the participants specified beforehand. The student tracking form is central in monitoring whether this goal has been achieved.

The test administrator plays a central role in all of these issues. Special consideration is therefore given to the training of the test administrators, ensuring that as little variation in the data as possible is caused by random or systematic variation in the activities of test administrators.

An important part of the testing situation relates to the relationship between test administrators and test participants. Therefore, any personal interaction between test administrators and students, either in the past or in the testing situation, counteracts the goal of collecting data in a consistent fashion across countries and participants. Strict objectivity of the test administrator, on the other hand, is instrumental in collecting data that reflect the level of literacy obtained, and that are not influenced by factors unrelated to literacy. The results based on these data are representative for the population under consideration.

Standard 8.1 All test sessions follow international procedures as specified in the PISA-D operations manuals, particularly the procedures that relate to:

- test session timing,
- maintaining test conditions,
- responding to students' questions,
- student tracking, and

- assigning assessment materials.

Standard 8.2 The relationship between Test Administrators and participating students must not compromise the credibility of the test session. In particular, the Test Administrator should not be the reading, mathematics, or science instructor of any student in the assessment sessions he or she will administer for PISA-D.

Note 8.1 Test Administrators should preferably not be school staff.

Quality control

- Session Report Forms
- PISA-D Quality Monitors (only for Main Survey) and their Data Collection Forms
- Main Survey Review Form

9. Training Support

Rationale: NPMs or their designees participated in a train-the-trainer session conducted by qualified contractor staff. This session facilitates standardisation of training delivery to test administrators, allows trainers to become familiar with PISA-D materials and procedures, and informs trainers of their responsibilities for overseeing the PISA-D testing.

Standard 9.1 Qualified contractor staff will conduct trainer training sessions with NPMs or designees on PISA-D materials and procedures to prepare them to train PISA-D test administrators.

Standard 9.2 NPMs or designees shall use the comprehensive training package developed by the contractors to train PISA-D test administrators.

Standard 9.3 All test administrator training sessions ideally should be scripted to ensure consistency of presentations across training sessions and across countries. Failure to do so could cause errors in data collection and make results less comparable.

Standard 9.4 In-person test administrator trainings should be conducted by the NPMs or designees, unless a suitable alternative is *agreed upon*.

Standard 9.5 PQMs need to successfully complete self-training materials, attend webinars to review and enhance the self-training, and attend the *test administrator* training, *unless otherwise agreed upon*.

Quality control

- Participation in trainer training sessions in standardised procedures by qualified contractor staff
- National Centre Quality Monitoring
- Main Survey Review Forms
- Standard training of PQMs and review of Data Collection Forms

10. Implementation of national options

Rationale: These standards serve to ensure that for students participating both in the international and the national survey, the national instruments do not affect the data used for the international comparisons. Data are therefore collected consistently across countries, and potential effects like test fatigue or learning effects from national test items are precluded.

Standard 10.1 Only *national options* that are *agreed upon* between the National Centre and the international contractors are implemented.

Standard 10.2 Any *national option* instruments that are not part of the core components of PISA-D are administered after all the test and questionnaire instruments of the core component of PISA-D have been administered to students that are part of the international PISA-D sample.

11. Security of the material

Rationale: The goal of the PISA-D assessment is to measure the literacy levels in the content domains. Prior familiarisation with the test materials, or training to the test, will heavily degrade the consistency and validity of the data. In the extreme case, the results would only reflect how well participants are able to memorise the test items. In order to be able to assess the competencies obtained during schooling rather than short-term learning success, and to make valid international comparisons, confidentiality is extremely important.

Standard 11.1 PISA-D materials designated as secure are kept confidential at all times. Secure materials include all test materials, data, and draft materials. In particular:

- no one other than approved project staff and participating students during the test session is able to access and view the test materials or Student Questionnaire responses,
- no one other than approved project staff will have access to secure PISA-D data and embargoed material, and
- formal confidentiality arrangements will be in place for all approved project staff.

Quality control

- Security arrangements as specified in the PISA-D operations manuals or agreed-upon variation
- National Centre Quality Monitoring
- PISA-D Quality Monitors (only for Main Survey) and Data Collection Forms
- Field Trial and Main Survey Review Forms

12. Quality Monitoring

Rationale: To obtain valid results from the assessment, the data collected have to be of high quality, that is, they have to be collected in a consistent, reliable, and valid fashion. This goal was

implemented first and foremost by the test administrators, who were seconded by the quality monitors. The quality monitors provided country-wide supervision of all data collection activities for the Main Survey.

Standard 12.1 PISA-D Main Survey test administration is monitored using site visits by trained independent quality monitors.

Standard 12.2 Fifteen site visits to observe test administration sessions are conducted in each PISA-D participating country/economy, and five site visits in each adjudicated region.

Standard 12.3 Test administration sessions that are the subject of a site visit are selected by the international contractors to be representative of a variety of schools in a country/economy.

Note 12.1 A failure to meet the Quality Monitoring standards in the Main Survey could lead to a significant lack of reliable and valid quality control information.

Note 12.2 The Quality Monitoring standards apply to the Main Survey but not to the Field Trial.

Note 12.3 The National Centre provides the international contractors the assistance required to implement the site visits effectively. This includes nominating sufficient qualified individuals to ensure that the required number of schools is observed.

Quality control

- The process of selecting the PISA-D Quality Monitor nominees
- PISA-D Quality Monitor (only for Main Survey) and their Data Collection Forms

13. Assembling and printing paper-based materials

Rationale: Variations in assembly and print quality may affect data quality. When the quality of paper and print is very poor, the performance of students is influenced not only by their levels of literacy, but also by the degree to which test materials are legible. To rule out this potential source of error, and to increase the consistency and precision of the data collection, paper and print quality samples were solicited from National Centres participating in paper-based components in their first cycle of participation.

Standard 13.1 All paper-based student assessment material will be centrally assembled by the international contractors and must be printed using the final print-ready file and *agreed-upon* paper and print quality.

Standard 13.2 The cover page of all PISA-D test paper-based materials used for students and schools must contain all information as specified by the PISA Governing Board.

Standard 13.3 The layout and pagination of all test paper-based material is the same as in the *source versions*, unless otherwise agreed upon.

Standard 13.4 The layout and formatting of the paper-based questionnaire material is equivalent to the source versions, with the exception of changes made necessary by national adaptations.

Note 13.1 For National Centres that have participated in previous cycles, PISA instruments used in previous cycles or from the Field Trial preceding the Main Survey that have been submitted to the previous international contractor can be used for the purpose of agreeing on printing quality where the National Centre indicates that printing and paper of the same standard will be used. Otherwise, National Centres will submit a sample of printed material to the international contractors for agreement, including the cover and selected items as specified in the PISA operations manuals.

Note 13.2 The cover page of all PISA-D PBA instruments used in schools should contain all information necessary to identify the material as being part of the data-collection process for PISA-D, and for checking whether the data collection follows the assessment design, i.e. whether the mapping of the student on the one hand, and test booklets and questionnaires, on the other, have been correctly established. The features of the cover page referred to in Standard 13.2 are specified in the PISA-D operations manuals.

Quality control

- Submitted sample materials or agreement that quality was similar to Field Trial versions
- Materials submitted to international contractors, as described in note 13.1 above
- Field Trial and Main Survey Review Forms

15. Response coding

Rationale: To ensure the comparability of the data, the responses from all test participants in all participating countries have to be coded following approved coding designs that are presented to both the Field Trial and the Main Survey. Therefore, all coding procedures have to be standardised, and coders have to complete training sessions to master this task.

Standard 15.1 Each *PISA-D participant* submits its data in a single complete database, unless otherwise *agreed upon*.

Standard 15.2 All *data* collected for PISA-D will be imported into a national database using the Data Management Expert (DME) data integration software provided by the international contractors following specifications in the corresponding operational manuals and international/national record layouts (codebooks). Data are submitted in the DME format.

Standard 15.3 Data for all *instruments* are submitted. This includes the assessment data, questionnaires data, and tracking data as described in the *PISA-D operations manuals*.

Standard 15.4 Unless *agreed upon*, all data are submitted without recoding any of the original response variables.

Standard 15.5 Each PISA-D participating country's database is submitted with full documentation as specified in the *PISA-D operations manuals*.

MANAGEMENT STANDARDS

16. Communication with the international contractors

Rationale: Given the tight schedule of the project, delays in communication between the National Centres and the international contractors should be minimised. Therefore, National Centres need continuous access to the various resources provided by the international contractors.

Standard 16.1 The international contractors ensure that qualified staff are available to respond in English to requests by the National Centres during all stages of the project. The qualified staff:

- Are authorised to respond to National Centre queries,
- Acknowledge receipt of National Centre queries within one working day,
- Respond to coder queries from National Centres within one working day,
- Respond to other queries from National Centres within five working days, or, if processing the query takes longer, give an indication of the amount of time required to respond to the query.

Standard 16.2 The National Centre ensures that qualified staff are available to respond to requests by the international contractors during all stages of the project. The qualified staff:

- Are authorised to respond to queries,
- Are able to communicate in English,
- Acknowledge receipt of queries within one working day,
- Respond to queries from international contractors within five working days, or, if processing the query takes longer, give an indication of the amount of time required to respond to the query.

Note 16.1 Response timelines and feedback schedules for the National Centres and the international contractor are further specified in the PISA-D operations manuals.

17. Notification of international and national options

Rationale: Given the tight timelines, the deadlines given in the following two standards enable the international contractors to progress with work on time.

Standard 17.1 National options are *agreed upon* with the international contractors before 1 December in the year preceding the Field Trial and before 1 November in the year preceding the Main Survey.

Standard 17.2 The National Centre notifies the OECD Secretariat of its intention to participate in specific international options three months prior to the start of the translation period.

18. Schedule for submission of materials

Rationale: To meet the requirements of the work programme, and to progress according to the timelines of the project, the international contractor needs to receive a number of materials on time.

Standard 18.1 An *agreed-upon Translation Plan* will be negotiated between each National Centre and the international contractors.

Standard 18.2 The *following* items are submitted to the international contractors in accordance with *agreed timelines*:

- the Translation Plan
- a print sample of booklets prior to final printing, for countries using the paper-based instruments (where this is required, see Standard 13.1 and Note 13.1),
- results from the national checking of adapted computer-based assessment materials and questionnaires,
- adaptations to school-level materials,
- sampling forms (see Standard 1),
- demographic tables,
- completed Field Trial and Main Survey Review Forms, and
- other documents as specified in the PISA-D operations manuals.

Standard 18.3 *Questionnaire* materials are submitted for linguistic verification only after all adaptations have been *agreed upon*.

Standard 18.4 All adaptations to those elements of the school-level materials that are required to be functionally equivalent to the source as specified in Standard 5.2, need to be *agreed upon*.

Quality control

- Agreed-upon Translation Plan
- International contractors' records from communications, forms, or documents
- Assessment materials submitted for linguistic verification with corresponding adaptation spreadsheets filled in by the National Centre

19. Management of data

Rationale: Consolidating and merging the national databases is a time-consuming and difficult task. To ensure the timely and efficient progress of the project, the international contractors need continuous access to national resources helping to rule out uncertainties and to resolve discrepancies. This standard aims to prevent substantial delays to the whole project which could result from a delay in processing the data of a small number of participating countries.

Standard 19.1 The timeline for submission of national databases to the international contractors is within eight weeks of the last day of testing for the Field Trial and within eight weeks of the last day of testing for the Main Survey, unless otherwise *agreed upon*.

Standard 19.2 National Centres execute data checking procedures as specified in the *PISA-D operations manuals* before submitting the database.

Standard 19.3 National Centres make a data manager available upon submission of the database. The data manager:

- is authorised to respond to international contractor data queries,
- is available for a three-month period immediately after the database is submitted unless otherwise *agreed upon*,
- is able to communicate in English,
- is able to respond to international contractor queries within three working days, and
- is able to resolve data discrepancies.

Standard 19.4 A complete set of PISA-D paper-based instruments as administered and including any *national options*, is forwarded to the international contractors on or before the first day of testing. The submission includes the following:

- hard copies of instruments, and
- electronic PDF copies of instruments

Standard 19.5 To enable the *PISA participant* to submit a single dataset, all instruments for all *additional adjudicated entities* will contain the same variables as the *primary adjudicated entity* of the *PISA participant*.

Note 19.1 Each participating country/economy will receive its own national micro-level PISA-D database (the “national database”), in electronic form and delivered as agreed upon a pre-specified timeline that varies based on their data submission. The national database will contain the complete set of responses from the students, school principals, and surveyed participants (*parents, teachers*) in that country/economy.

Each participating country/economy has access to and can publish its own data after a date that is established by the PISA Governing Board for the publication of the initial OECD publication of the survey results (the “initial international OECD publication”).

The OECD Secretariat will not release national data to other countries/economies until participating countries/economies have been given an opportunity to review and comment on their own national data and until the release of such data has been approved by the national authorities.

A deadline and procedures for withdrawing countries/economies’ national data from the international micro-level PISA-D database (the “international database”) will be decided upon by the PISA Governing Board. Countries/economies can withdraw data only prior to obtaining access to data from other countries/economies. Withdrawn data will not be made available to other countries/economies.

The PISA Governing Board will discuss with participating countries/economies whose data manifests technical anomalies as to whether the data concerned can be included in the international database. The decision of the PISA Governing Board will be final. Participating countries/economies may, however, continue to use data that are excluded from the international database at the national level.

The international contractors will then compile the international database, which will comprise the complete set of national PISA-D databases, except those data elements that have been withdrawn by participating countries/economies or by the PISA Governing Board at the previous stage. The international database will remain confidential until the date on which the initial international OECD publication is released.

National data from all participating countries/economies represented in the international database will be made available to all participating countries/economies from the date on which the initial international OECD publication is released.

After release of the initial international OECD publication, the international database will be made publicly available on a cost-free basis, through the OECD Secretariat. The database may not be offered for sale.

The international database will form the basis for OECD indicator reports and publications.

The international contractors will have no ownership of instruments or data nor any rights of publication and will be subject to the confidentiality terms set in this agreement.

The OECD establishes rules to ensure adherence to the above procedure and to the continued confidentiality of the PISA-D data and materials until the agreed release dates. These include confidentiality agreements with all individuals that have access to the PISA-D material prior to its release.

As guardian of the process and producer of the international database, the OECD will hold copyright in the database and in all original material used to develop, or be included in, the PISA-D Field Trial and PISA-D Main Survey (among them the assessment materials, school-level materials, and coding guides) in any language and format.

Quality control

- International contractors' records

20. Archiving of materials

Rationale: The international contractors maintained an electronic archive. This provided an overview of all materials used and ensures continuity of materials available in participating countries *across PISA survey cycles*, therefore building upon the knowledge gained nationally in the course of the PISA-D cycles. This also ensured that the international contractors had the relevant materials available during data cleaning, when they were first required.

Standard 20.1 The international contractors will maintain a permanent electronic archive of all assessment materials, school-level materials, and coding guides, including all national versions.

Standard 20.2 The National Project Manager must submit one copy of each of the following translated and adapted Main Survey materials to the international contractors:

- a hard copy of all administered Test Instruments, including *national options* (paper-based countries only);
- a hard copy of all administered Questionnaires, including *national options* (paper-based countries only);
- electronic versions of the school-level materials; and
- electronic versions of the Coding Guides.

Standard 20.3 Unless otherwise requested, National Centres will archive all Field Trial materials until the beginning of the Main Survey, and all Main Survey materials until the publication of the international report. Materials to be archived include:

- all respondents' paper-based test booklets and questionnaires (PBA countries) or USB drives containing completed SDS and all associated data (CBA countries),
- sampling forms,
- student lists,
- student tracking instruments, and

- all data submitted to the international contractors.

After completion of a survey the National Centre will transfer this archive to the international contractors who will compile the national archives from all participants and transfer them to OECD after completion of the Main Survey.

NATIONAL INVOLVEMENT STANDARDS

21. National feedback

Rationale: National feedback in areas such as test development is important in maintaining the dynamic and collaborative nature of PISA-D. National feedback ensures that instruments achieve cross-national, cross-cultural, and cross-linguistic validity. It also promotes the inclusion of the interests and involvement of national stakeholders.

Standard 21.1 National Centres develop appropriate mechanisms in order to promote participation, effective implementation, and dissemination of results amongst all relevant national stakeholders.

Standard 21.2 National Centres provide feedback to the international contractors on the development of instruments, domain frameworks, the adaptation of instruments, and other domain-related matters that represent the perspectives of the relevant national stakeholders.

Note 21.1 As a guideline, feedback might be sought from the following relevant stakeholders: policy makers, curriculum developers, domain experts, test developers, linguistic experts and experienced teachers.

Quality control

- National Centre quality monitoring
- List of committees and groups of stakeholders
- Membership records of representative groups and/or committees
- Meeting records of representative groups and/or committees

22. Meeting attendance

Rationale: Attendance at National Project Managers and training meetings is required as these represent a key component of participating in PISA-D. Important information was shared and discussed and training in data management, sampling, computer systems, and coding was conducted at these international meetings. These also allow for individual consultation and communication with contractors, which is often very helpful.

Standard 22.1 Representatives from each National Centre are required to attend all PISA-D international meetings including National Project Manager meetings, coder training and data management training. Up to eight international meetings are planned per cycle.

Standard 22.2 Representatives from each National Centre who attend international meetings must be able to work and communicate in English.

Note 22.1 The length of these meetings vary from 3 to 5 days.

Quality control

- Meeting attendance records

DEFINITIONS

Additional Adjudicated Entities - entities in addition to the first and primary entity managed by a PISA-D participant, where a PISA participant manages more than one adjudicated entity.

Adjudicated Entity – a country, geographic region, or similarly defined population, for which the international contractors fully implement Quality Assurance and Quality Control mechanisms and endorse, or otherwise, the publication of separate PISA results.

Agreed procedures – procedures that are specified in the PISA-D operations manuals, or variations that are agreed upon between the National Project Manager and the international contractors.

Agreed timelines – timelines that are specified in the PISA-D operations manuals, or variations that are agreed upon between the National Project Manager and the international contractors.

Agreed-upon variations and definitions agreed upon between the National Project Manager and the international contractors.

International contractors website – The PISA Portal – PISA 2018 project website – can be accessed through the following address: <http://pisa.ets.org/portal>. This website contains the source versions of instruments, manuals and other documents and information relating to National Centres.

International Option - optional additional international instruments or procedures designed and fully supported by the international contractors.

MyMWestra – software developed specifically for the PISA-D project. The software assists with sampling, student tracking, and data submission practices that meet the PISA-D technical standards.

National Centre Quality Monitoring – the procedures by which the international contractors monitor the quality of all aspects of the implementation of the survey by a National Centre.

National Option – A national option occurs if:

- a) National Centre administers any additional instrumentation, for example a test or questionnaire, to schools or students that are part of the PISA-D international sample.

Note that in the case of adding items to the questionnaires, an addition of five or more items to either the school questionnaire or the student questionnaire is regarded as a national option.

OR

- b) National Centre administers any PISA-D international instrumentation to any students or schools that are not part of an international PISA-D sample (age-based or grade-based) and therefore will not be included in the respective PISA-D international database.

PISA Defined Target Population – all PISA-Eligible students in the schools that are listed on the school sampling frame. That is, the PISA Desired Target Population minus exclusions.

PISA Desired Target Population – the PISA Target Population defined for a specific adjudicated entity. It provides the most exhaustive coverage of PISA-Eligible students in the participating economy as is feasible.

PISA-Eligible Students – students who are in the PISA Target Population.

PISA Target Population – students aged between 15 years and 3 (completed) months and 16 years and 2 (completed) months at the beginning of the testing period, attending educational institutions located within the adjudicated entity, and in grade 7 or higher. The age range of the population may vary up to one month, either older or younger, but the age range must remain 12 months in length. That is, the population can be as young as between 15 years and 2 (completed) months and 16 years and 1 (completed) month at the beginning of the testing period; or as old as between 15 years and 4 (completed) months and 16 years and 3 (completed) months at the beginning of the testing period.

PISA-D Operations Manuals – all manuals provided by the international contractors. The preparation of the PISA operations manuals are carried out by the international contractors and describe procedures developed by the international contractors. The manuals are prepared following consultation with participating countries/economies, the OECD Secretariat, the Technical Advisory Group and other stakeholders.

PISA-D Participant – an administration centre, commonly called a National Centre that is managed by a person or persons, commonly called a National Project Manager(s), who is/are responsible for administering PISA in one or more adjudicated entities. The National Project Manager(s) must be authorised to communicate with the international contractor on all operational matters relating to the adjudicated entities for which the National Project Manager is responsible.

PISA-D Quality Monitor – a person nominated by the National Project Manager and employed by the international contractors to monitor test administration quality in an adjudicated entity, if applicable.

PISA-D Quality Monitor Report – a report produced by the international contractor overseeing the PISA-D Quality Monitor process. The report contains a summary of the feedback collected by the PISA-D Quality Monitors.

PISA-D SharePoint Site – The PISA-D project website, which can be accessed through the following address: <https://etsorg1.sharepoint.com/sites/ext/PISA-D/SitePages/Home.aspx>. This website contains the source versions of instruments, manuals and other documents and information relating to National Centres.

School Associate – a person at a school who serves as both a School Co-ordinator and Test Administrator.

School Co-ordinator – a person at a school who acts as a liaison between the school and the National Centre to prepare for the assessment in the school.

School-level Exclusions – contractors’ approved exclusion of schools from the sampling frame because:

- of geographical inaccessibility (but not part of a region that is omitted from the PISA Desired Target Population),
- administration of the PISA-D assessment within the school would not be feasible,
- all students in the school would be within-school exclusions, or
- of other reasons as agreed upon.

School-level Materials – the key materials include:

- School Co-ordinator Manual and Test Administrator Manual (or School Associate Manual)
- Test Administrator script
- Key forms – Student Tracking Form, Session Attendance Form, and Session Report Form

Source or Base Versions – documents provided in English (and, for some documents, French and Spanish) by the international contractors.

Target Cluster Size – the number of students that are to be sampled from schools where not all students are to be included in the sample.

Test Administrator – a person who is trained by the National Centre to administer the PISA-D test in schools. This person may be a Test Administrator or a School Associate (a School Co-ordinator who also has the role of a Test Administrator).

Testing Period – the period of time during which data is collected in an adjudicated entity.

Translation Plan – documentation of all the processes that are intended to be used for all activities related to translation and languages.

Within-school exclusions – exclusion of students from potential assessment because of one of the following:

- They are functionally disabled in such a way that they cannot take the PISA-D test. Functionally disabled students are those with a moderate to severe permanent physical disability.
- They have a cognitive, behavioural, or emotional disability confirmed by qualified staff, meaning they cannot take the PISA-D test. These are students who are cognitively, behaviourally, or emotionally unable to follow even the general instructions of the assessment.
- They have insufficient assessment language experience to take the PISA-D test. Students who have insufficient assessment language experience are those who meet all the following three criteria:
 - they are not native speakers of the assessment language,
 - they have limited proficiency in the assessment language, and
 - they have received less than one year of instruction in the assessment language.
- There are no materials available in the language in which the student is taught.
- They cannot be assessed for some other reason as agreed upon.