



ANNEX F – TECHNICAL STANDARDS FOR PISA 2012

INTRODUCTION

The purpose of this annex is to list the set of standards upon which the PISA 2012 data collection activities will be based, as was the case for previous PISA assessments [EDU/PISA/GB(2007)4/REV1]. In following the procedures specified in the standards, the partners involved in the data collection activities contribute to creating an international dataset of a quality that allows for valid cross-national inferences to be made.

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. A comparable sample of the 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 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 un-related 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, e.g. 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, test materials and tasks, etc. should be selected 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 in the target population at large. Thus, collecting data from a representative sample of the population, for example, will lead to results that accurately reflect the level of literacy of fifteen-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, as well as the sheer number of participating countries, preclude the option of developing and monitoring local solutions to be harmonised at a later stage in the project. Therefore, the standards specify one clear-cut path along which data collection and data submission should progress.

This document strives to establish a collective agreement of mutual accountability among countries, and of the international contractor towards the countries. This document details each standard, its rationale, and the quality assurance data that need to be collected to demonstrate that the standard has been met.

Where standards have been fully met, data will be recommended for inclusion in the PISA 2012 dataset. Where standards have not been fully met, an adjudication process will determine the extent to which the quality and international comparability of the data have been affected. The result of data adjudication will determine whether the data will be recommended for inclusion in the PISA 2012 dataset.

Since attaining the various standards is cumulative and potentially interactive (i.e. not attaining standard X is NOT the same as not attaining standards X, Y and Z), in principle each dataset should be evaluated against all standards jointly. Also, it is possible that countries' proposed plans for implementation are 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 which impact on print and data management quality). Therefore, the final evaluation of standards needs to be made with respect to the data as submitted since this is the definitive indication of what may appear in the released international dataset.

If any issues with attaining standards are identified, the International Project Director initiates communication with the National Centre as soon as possible. Priority in communication rectifies the identified issues.

The PISA standards act as a benchmark of best practice. As such, the standards are designed to assist National Centres and international contractor 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 contractor are defined in the *PISA Operations Manuals*.

As specified in the contracts for the implementation of the fifth cycle of the OECD Programme for International Student Assessment, the international contractor takes responsibility for developing and implementing procedures for assuring data quality. Therefore, the international contractor mediates, and monitors the countries' activities specified in this document, while the international contractor's adherence to the standards is monitored by the participating countries via the OECD Secretariat.

Where the Technical Standards stipulate that variations from the standards require agreement between participating countries and the Consortium, National Project Managers are asked to initiate the process of negotiation and to undertake everything possible to facilitate an agreement. Where agreement between National Project Managers and the Consortium cannot be reached, the OECD will adjudicate and resolve the issues. The OECD will also adjudicate any issues resulting from non-compliance with the Technical Standards that cannot be resolved between participating countries and the Consortium.

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 quality of the data or the assurance of that quality. These standards have been endorsed by the Technical Advisory Group and wherever proportions or quantities are specified (for example, response rates), these have reached through examination of research undertaken or 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 operational objectives are met in a timely and coordinated manner.
- **National Involvement Standards** reflect the expectations set out in the PISA 2012 Terms of Reference that the content of the PISA tests is established in consultation with national representatives with international content expertise. In particular, these standards 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 THE DOCUMENT

The standards are grouped into sections that relate to specific tasks in the PISA 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 Assurance** measures that will be 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 both *PISA participants* and *additional adjudicated entities*. The PISA Governing Board will approve the list of *adjudicated entities* to be included in a PISA cycle.

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 will not be confounded by potential age effects. Furthermore, to be able to draw conclusions that are valid for the entire population of fifteen-year-old students, a representative sample shall be 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 are specified.



- Standard 1.1** The *PISA Desired Target Population* is agreed upon through negotiation between the National Project Manager and the international contractor, within the constraints imposed by the definition of the *PISA Target Population*.
- Standard 1.2** Unless otherwise agreed upon only *PISA-Eligible students* participate in the test.
- Standard 1.3** Unless otherwise agreed upon, the *testing period*:
- is no longer than six consecutive weeks in duration;
 - 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** 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 1.6** 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 1.7** The student sample size is a minimum of 4 500 assessed students for *PISA participants* and 1 500 assessed students for *additional adjudicated entities*, or the entire *PISA Defined Target Population* where the *PISA Defined Target Population* is below 4 500 and 1 500 respectively.
- Standard 1.8** The school sample size is a minimum of 150 schools for *PISA participants*, and 50 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.
- Standard 1.9** The school response rate is at least 85% of sampled schools. If a response rate is below 85% then an acceptable response rate can still be achieved through *agreed upon* use of replacement schools.
- Standard 1.10** The student response rate is at least 80% of all sampled students across responding schools.

Note 1.1 The Target Population and Sampling standard apply to the Main Survey but not the Field Trial.

Note 1.2 Data from schools where the student response rate is greater than 25 % will be included in the PISA 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 students respond.

Note 1.4 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 would be not feasible within the school, and other *agreed upon* reasons and that 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 1.5 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 a *target cluster size* of 35 *PISA-Eligible students* which *upon agreement* can be increased, or reduced to a number not less than 20.

Quality assurance

- Sampling procedures as specified in the *PISA Operations Manuals*
- School sample drawn by international contractor (or if drawn by the National Centre, then verified by the international contractor)
- Student sample drawn through *KeyQuest* (or if drawn by other means, then verified by the international contractor)
- Sampling forms submitted to the international contractor
- Main Survey Review Quality Assurance Survey

2. Language of testing

Rationale: Using the language of instruction will ensure 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 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 in the major domain (mathematics) of the test.

- 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 test booklets 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 booklets with more than one language in the Main Survey.
- In all cases the choice of test language(s) in the test booklets 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 contractor to make detailed analyses of the items so that only suitable ones are included in the Main Survey.

Standard 3.1. *PISA participants* participating in the PISA 2012 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 assessment languages that apply to between 5 and 50% of the target population, the Field Trial student sample should be a minimum of 100 students per item.
- For languages that apply to more than 50% of the target population, the Field Trial student sample should be a minimum of 200 students per item.
- For *additional adjudicated entities*, where the assessment language applies to between 5 and 100% 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 Field Trial participation standard for assessment languages applicable to between 5 and 50% of the Target Population can be varied if *agreed upon*, with such agreement subject to the principle that the absence of a Field Trial for that language would not affect the Main Survey and the principle that the assessment language version is trialled in another *adjudicated entity* where the assessment language applies to more than 50% of the Target Population.

Note 3.3 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 3.4 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 manuals

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 the 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 for the assessment instruments that are used within a PISA cycle: To validly compare performance across countries, all assessment instruments have to be as similar as possible. In fact, it is of utmost importance to provide equivalent information for 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, i.e. the Test Administrators, School Co-ordinators, and school associates, should receive the same information in all participating countries.



- Standard 4.1** Test items used for linking are administered unchanged from their previous administration.
- Standard 4.2** All test instruments are psychometrically equivalent to the *source versions*. *Agreed upon* adaptations to the local context are made if needed.
- Standard 4.3** The questionnaire instruments are equivalent to the *source versions*. *Agreed upon* adaptations to the local context are made if needed.
- Standard 4.4** The Test Administrator Manual and the School Co-ordinator Manual (or the School Associate Manual) are equivalent to the *source versions*. *Agreed upon* adaptations to the local context are made if needed.

Note 4.1 The quality assurance requirements for this standard apply to instruments that are in an assessment language used as a language of instruction for more than 5% of the Target Population.

Quality assurance

- *Agreed upon* Manual Adaptation Spreadsheet (MAS) and Questionnaire Adaptation Spreadsheet (QAS)
- Test Adaptation Spreadsheet (TAS) and Booklet Adaptation Spreadsheet (BAS) in which adaptations to assessment units and common booklet parts are documented. Adaptations will be checked for compliance with the PISA Translation and Adaptation Guidelines by international verifiers, and the verifiers' recommendations will be vetted by the Consortium referee.
- Verifier Reports (statistics generated by the TAS with a short qualitative report)
- Final Optical Check (FOC) Report (test booklets and questionnaires only) including key correction check
- Post-FOC review
- Field Trial and Main Survey Review Quality Assurance Surveys
- Item and scale statistics

5. Translation of tests, questionnaires and manuals

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 will be 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 test 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*, the following documents are translated/adapted into the assessment language to make them linguistically equivalent to the international *source versions*.
- The Test Administrator (or School Associate) Manual
 - The School Co-ordinator (or School Associate) Manual
- In the case of the manuals, only *specified parts* are made linguistically equivalent.

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.

Note 5.2 The "specified parts" of manuals referred to in Standard 5.2 for which checking of the linguistic equivalence to the source versions would be undertaken are the following:

- The criteria for student eligibility
- The number of students to be sampled from each school
- The definitions, codes and instructions related to the coding of the Student Tracking Form, including examples to illustrate these codes
- The General Directions as well as instructions relating to the timing of sessions
- The Session Report Form completed by the Test Administrator for each testing session, which records session and timing information.



Quality assurance

- *Agreed upon Translation Plan* developed in accordance with the specifications in the *PISA Operations Manuals* where the *Translation Plan* would require double translation by independent translators from two *source versions*.
- *Agreed upon Questionnaire Adaptation Spreadsheet (QAS)*
- *Test Adaptation Spreadsheet (TAS)* and *Booklet Adaptation Spreadsheet (BAS)* in which adaptations to assessment units and common booklet parts are documented. Adaptations will be checked for compliance with the PISA Translation and Adaptation Guidelines by international verifiers, and the verifiers' recommendations will be vetted by the Consortium referee.
- *Verifier Reports* (statistics generated by the TAS and short qualitative report)
- *Final Optical Check report* (test booklets and questionnaires only)
- *Submitted test booklets as used in the study*
- *Field Trial and Main Survey Review Quality Assurance Surveys*
- *Item and scale statistics*

6. 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 rulers 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 *PISA Operations Manuals*. 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 will the students participating in the study mirror the population of fifteen-year-old students in the country. The statistical analysis will take 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 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 un-related to literacy. The results based on these data will be representative for the population under consideration.

Standard 6.1 All test sessions follow international procedures as specified in the *PISA Operations Manuals*, particularly the procedures that are:

- relating to test session timing;
- for maintaining test conditions;
- for student tracking; and
- for assigning booklets.

Standard 6.2 Test Administrators are trained in person unless a suitable alternative is *agreed upon*.

Standard 6.3 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.

Note 6.1 Test Administrators should preferably not be school staff.

Note 6.2 Preferred training procedures for Test Administrators are described in the *PISA Operations Manuals*.



Quality assurance

- Test Administrator's Test Session Report Forms
- PISA Quality Monitors
- Main Survey Review Quality Assurance Survey

7. 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 will 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 7.1** Only *national options* that are *agreed upon* between the National Centre and the international contractor are implemented.
- Standard 7.2** Any *national option* instruments that are not part of the core component of PISA are administered after all the test and questionnaire instruments of the core component of PISA have been administered to students that are part of the international PISA sample.

8. Security of the material

Rationale: The goal of the PISA 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 8.1** PISA 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 material;
 - no-one other than approved project staff will have access to secure PISA data and embargoed material; and
 - formal confidentiality arrangements will be in place for all approved project staff.

Quality assurance

- Security arrangements as specified in the *PISA Operations Manuals* or *agreed upon* variation
- National Centre Quality Monitoring
- Field Trial and Main Survey Review Quality Assurance Surveys

9. Quality monitoring

Rationale: To obtain valid results from the assessment, the data collected have to be of high quality, i.e. they have to be collected in a consistent, reliable and valid fashion. This goal is implemented first and foremost by the test administrators, who are seconded by the quality monitors. The quality monitors provide country-wide supervision of all data-collection activities.

- Standard 9.1** PISA test administration is monitored using site visits by trained independent quality monitors.
- Standard 9.2** An agreed number of site visits to observe Test Administration sessions are conducted in each PISA participating country/economy.
- Standard 9.3** Test Administration sessions that are the subject of a site visit are randomly selected.

Note 9.1 A failure to meet the Quality Monitoring Standard in the Main Survey will lead to a significant lack of quality assurance data for other standards.

Note 9.2 The Quality Monitoring standards apply to the Main Survey but not to the Field Trial.

Note 9.3 The National Centre provides the international contractor the assistance required to implement the site visits effectively.



Quality assurance

- Curricula Vitae of the PISA Quality Monitor nominees forwarded by the National Project Manager to the international contractor
- PISA Quality Monitor Reports
- National Centre Quality Monitor Report

10. Printing of material

Rationale: Variations in 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 are solicited from National Centres in their first cycle of participation.

- | | |
|----------------------|--|
| Standard 10.1 | All student assessment material is printed using an agreed upon paper and print quality. |
| Standard 10.2 | The cover page of all PISA assessment instruments used in schools contains all information as specified by the PISA Governing Board. |
| Standard 10.3 | The layout and pagination of all test material is the same as in the <i>source versions</i> , unless otherwise agreed upon. |
| Standard 10.4 | The layout and formatting of the questionnaire material is equivalent to the <i>source versions</i> . |

Note 10.1 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 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 contractor for agreement, including the cover and selected items as specified in the *PISA operations manuals*.

Note 10.2 The cover page of all PISA assessment instruments used in schools should contain all information necessary to identify the material as being part of the data-collection process for PISA, 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 10.2 are specified in the *PISA Operations Manuals*.

Quality assurance

- Submitted sample or agreement that quality will be similar to previous cycle or Field Trial versions
- Booklets submitted to international contractor to meet Standard 17.4
- Booklets submitted for The International Coding Review (ICR) (Main Survey only)
- Field Trial and Main Survey Review Quality Assurance Surveys

11. Response coding¹

Rationale: To ensure the comparability of the data, the responses from all test participants in all participating countries have to be coded following one single coding scheme. Therefore, all coding procedures have to be standardised, and coders have to complete training sessions to master this task.

- | | |
|----------------------|---|
| Standard 11.1 | The coding scheme described in the coding guide in the distributed items is implemented according to instructions from the international contractor's item developers. |
| Standard 11.2 | Representatives from each National Centre attend the international PISA coder training session for both the Field Trial and the Main Survey. |
| Standard 11.3 | Both the single and multiple coding procedures as specified in the <i>PISA Operations Manuals</i> (see Note 11.1), or an <i>agreed upon</i> variation thereof, are implemented. |
| Standard 11.4 | Coders are recruited and trained following <i>agreed procedures</i> . |

Note 11.1 Preferred procedures for recruiting and training coders are outlined in the *PISA Operations Manuals*

Note 11.2 The optimum number of Coder Training session participants would depend on factors such as the expertise of National Centre staff, and resource availability.

.....
1. The terms coding, coders and codes are used instead of other terms such as marking, markers, marks, rating and raters.



Quality assurance

- Indices of inter-coder agreement
- International Coding Review (ICR)
- Coding of control scripts from previous cycle's ICR
- Field Trial and Main Survey Review Quality Assurance Surveys

12. Data submission

Rationale: The timely progression of the project, within the tight timelines given depends on the quick and efficient submission of all collected data. Therefore, one single data submission format is proposed, and countries are asked to submit only one database to the international contractor. Furthermore, to avoid potential errors when consolidating the national databases, any changes in format that were implemented subsequent to the general agreement have to be announced.

Standard 12.1	Each <i>PISA participant</i> submits its data in a single database, unless otherwise <i>agreed upon</i> .
Standard 12.2	Data are submitted in the <i>KeyQuest</i> format.
Standard 12.3	Data for all instruments are submitted. This includes the test data, questionnaire data, and tracking data as described in the <i>PISA Operations Manuals</i> .
Standard 12.4	Unless <i>agreed upon</i> , all data are submitted without recoding any of the original response variables.
Standard 12.5	Each <i>PISA participating country's</i> database is submitted with full documentation as specified in the <i>PISA Operations Manuals</i> .

MANAGEMENT STANDARDS

13. Communication with the international contractor

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

Standard 13.1	The international contractors ensure that qualified staff are available to respond to requests by the National Centres during all stages of the project. The qualified staff: <ul style="list-style-type: none"> ▪ 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; and ▪ 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.
----------------------	--

Note 13.1 Response timelines and feedback schedules for the National Centres and the international contractor are further specified in the *PISA Operations Manuals*.

14. Notification of international and national options

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

Standard 14.1	<i>National options</i> are <i>agreed upon</i> before 1 December in the year preceding the Field Trial and before 1 December in the year preceding the Main Survey (Standard 7.1).
Standard 14.2	The National Centre notifies the international contractor of its intention to participate in specific international options before 1 December in the year preceding the Field Trial.



15. 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 will need to receive a number of materials on time.

- Standard 15.1** An *agreed upon Translation Plan* and *Preferred Verification Schedule* will be negotiated between each National Centre and the international contractor.
- Standard 15.2** The following items are submitted to the international contractor in accordance with *agreed timelines*:
- the Translation Plan and Preferred Verification Schedule;
 - a print sample of booklets prior to final printing (where this is required, see Standard 10.1 and Note 10.1);
 - sampling forms (see Standard 1);
 - Demographic Tables;
 - Field Trial and Main Survey Reviews; and
 - other documents as specified in the *PISA Operations Manuals*.
- Standard 15.3** Questionnaire materials are submitted for linguistic verification only after all adaptations have been *agreed upon*.
- Standard 15.4** Those elements of the Test Administrator and School Co-ordinator (or School Associate) manuals requiring verification, including linguistic verification as specified in Standard 5.2, are submitted only after all adaptations have been *agreed upon*.

Quality assurance

- Agreed upon Translation Plan and Preferred Verification Schedule
- International contractor records
- Test materials are submitted for linguistic verification with corresponding adaptation spreadsheets filled in by the National Centre

16. Drawing samples

Rationale: The mode of drawing the samples used in the study is crucial to data quality. The goal of the project is to collect data that are representative for the population at large. To reach this goal, the sampling procedures have to follow established scientific rules. Furthermore, the comparability of the data across countries is guaranteed if the same procedure is used for all national samples. If different sampling procedures are used, then the equivalence of the sampling quality has to be determined.

- Standard 16.1** For efficient and effective quality assurance provision, unless otherwise *agreed upon*, the international contractor will draw the school sample for the Main Survey.
- Agreement with the international contractor will be subject to the principle that the sampling methods used are scientifically valid and consistent with PISA's documented sampling methods. Where a *PISA 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.
- Standard 16.2** For efficient and effective quality assurance provision, unless otherwise *agreed upon*, the National Centre will use *KeyQuest* to draw the student sample, using the list of eligible students provided for each school.
- Where a *PISA participating country* chooses not to use *KeyQuest* 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.

Note 16.1 Any costs associated with verifying a school sample taken by the National Centre, or a student sample selected other than by using *KeyQuest* will be borne by the National Centre.



17. 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 contractor needs 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 17.1** The timeline for submission of national databases to the international contractor is within eight weeks of the last day of testing for the Field Trial and within twelve weeks of the last day of testing for the Main Survey, unless otherwise *agreed upon*.
- Standard 17.2** National Centres execute data checking procedures as specified in the *PISA Operation Manuals* before submitting the database.
- Standard 17.3** 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 respond to international contractor queries within three working days; and
 - is able to resolve data discrepancies.
- Standard 17.4** A complete set of PISA instruments as administered and including any *national options*, is forwarded to the international contractor on or before the first day of testing. The submission includes the following:
- hard copies of instruments; and
 - electronic PDF copies of instruments
- Standard 17.5** 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 17.1 Each participating country/economy will receive its own national micro-level PISA database (the “national database”), in electronic form as soon as it has been processed from the international contractor for PISA. The national database will contain the complete set of responses from the students, parents, school principals and surveyed participants 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 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 OECD Secretariat will then compile the international database, which will comprise the complete set of national PISA 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 contractor for PISA 2012 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 data and materials until the agreed release dates. These include confidentiality agreements with all individuals that have access to the PISA 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 Field Trial and PISA Main Survey (among them the assessment materials, field manuals, and coding guides) in any language and format.

Quality assurance

- International contractor Records



18. Archiving of materials

Rationale: The international contractor will maintain an electronic archive. This will provide an overview of all materials used and ensure continuity of materials available in participating countries across PISA survey cycles, therefore building upon the knowledge gained nationally in the course of the PISA cycles. This will also ensure that the international contractor has the relevant materials available during data cleaning, when they are first required.

Standard 18.1 The international contractor will maintain a permanent electronic archive of all assessment materials, field manuals and coding guides. To facilitate this, the National Project Manager submits one copy of each of the following translated and adapted Main Survey materials to the international contractor in the source version software format:

- all administered Test Instruments, including *national options*;
- all administered Questionnaires, including *national options*;
- Test Administrator, School Co-ordinator and School Associate manuals; and
- Coding Guides.

Standard 18.2 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' test booklets and questionnaires;
- sampling forms;
- student lists;
- student tracking instruments; and
- all data submitted to the international contractor.

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

NATIONAL INVOLVEMENT STANDARDS

19. National feedback

National feedback in areas such as test development is important in maintaining the dynamic and collaborative nature of PISA. 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 19.1 National Centres develop appropriate mechanisms in order to promote participation, effective implementation, and dissemination of results amongst all relevant national stakeholders.

Standard 19.2 National Centres provide feedback to the international contractor on the development of instruments, domain frameworks, the adaptation of instruments, and other domain related matters that represents the perspectives of the relevant national stakeholders.

Note 19.1 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 assurance

- National Centre Quality Monitoring
- Documented strategies
- List of committees and groups
- Membership records of representative groups and/or committees
- Meeting records of representative groups and/or committees



DEFINITIONS

Additional Adjudicated Entities – entities in addition to the first and primary entity managed by a *PISA 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 contractor fully implements quality assurance and quality control mechanisms and endorses, or otherwise, the publication of separate PISA results.

Agreed procedures – procedures that are specified in the *PISA Operations Manuals*, or variations that are *agreed upon* between the National Project Manager and the international contractor.

Agreed timelines – timelines that are specified in the *PISA Operations Manuals*, or variations that are *agreed upon* between the National Project Manager and the international contractor.

Agreed upon – variations and definitions agreed upon between the National Project Manager and the international contractor. Agreed upon variations are available to National Project Managers on their National Centre webpage on the *international contractor Website*.

International contractor website – website with address *http://mypisa.acer.edu.au*. This website contains the *source versions* of instruments, manuals and other documents and information relating to National Centres. These materials are also available from *www.oecd.org/pisa*.

International Coding Review – a quality assurance exercise that requires National Centres to send a sample of student test booklets to the international contractor. The booklets required for the quality assurance study will be identified by the international contractor after the National Centre's data has been submitted. The number of booklets to be submitted by each *PISA participating country/economy* will depend on the number of languages of assessment, the number of adjudicated entities, and the number of coding centres used.

International Option – optional additional international instruments or procedures designed and fully supported by the international contractor.

KeyQuest – software developed by the international contractor specifically for the PISA project. The software assists with sampling, student tracking and data submission practices that meet the PISA 2012 technical standards.

National Centre Quality Monitor – an international contractor representative who visits a National Centre in the month preceding the Main Survey to train *PISA Quality Monitors* and conduct a scheduled interview with the National Project Manager.

National Option – A *national option* occurs if:

i) A National Centre administers any additional instrumentation, for example a test or questionnaire, to schools or students that are part of the PISA 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

ii) A National Centre administers any PISA international instrumentation to any students or schools that are not part of an international PISA sample (age-based or grade-based) and therefore will not be included in the respective PISA 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 *adjudicated entity* as is feasible.



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

PISA Operations Manuals – manuals provided by the international contractor, that is the following:

- National Project Manager’s Manual;
- Test Administrator Manual;
- School Coordinator Manual;
- School Associate Manual;
- School Sampling Preparations Manual;
- Data Management Manual; and
- all other key documents referenced within the National Project Manager’s manual.

The preparation of the *PISA Operations Manuals* will be carried out by the international contractor and will describe procedures developed by the international contractor. The manuals will be prepared following consultation with participating countries/economies, the OECD Secretariat, the Technical Advisory Group and other stakeholders.

PISA Participant – an administration centre, commonly called a National Centre that is managed by a person, commonly called a National Project Manager, who is responsible for administering PISA in an *adjudicated entity* and in zero or more *additional adjudicated entities*. The National Project Manager 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 Quality Monitor – a person nominated by the National Project Manager and employed by the international contractor to monitor Test Administration quality in an adjudicated entity.

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*.

Preferred Verification Schedule – a schedule that provides a timeline for the submission of material relating to the adaptation of instruments and the submission of instruments for linguistic verification including the Final Optical Check. This schedule can be found in the PISA National Project Manager’s Manual.

School Level Exclusions – 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*);
- of an extremely small size;
- administration of the PISA 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*.

Source Versions – documents provided in English and French by the international contractor.

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.

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 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 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 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.
- They cannot be assessed for some other reason as *agreed upon*.