



Annex A

PISA 2012 TECHNICAL BACKGROUND

All figures and tables in Annex A are available on line

Annex A1: Construction of mathematics scales and indices from the student, school and parent context questionnaires
<http://dx.doi.org/10.1787/888932937073>

Annex A2: The PISA target population, the PISA samples and the definition of schools
<http://dx.doi.org/10.1787/888932937092>

Annex A3: Technical notes on analyses in this volume

Annex A4: Quality assurance

Annex A5: Technical details of trends analyses
<http://dx.doi.org/10.1787/888932960500>

Annex A6: Anchoring vignettes in the PISA 2012 Student Questionnaire

Notes regarding Cyprus

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

Note by all the European Union Member States of the OECD and the European Union: The Republic of Cyprus is recognised by all members of the United Nations with the exception of Turkey. The information in this document relates to the area under the effective control of the Government of the Republic of Cyprus.

A note regarding Israel

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



ANNEX A1

CONSTRUCTION OF MATHEMATICS SCALES AND INDICES FROM THE STUDENT, SCHOOL AND PARENT CONTEXT QUESTIONNAIRES

How the PISA 2012 mathematics assessments were designed, analysed and scaled

The development of the PISA 2012 mathematics tasks was co-ordinated by an international consortium of educational research institutions contracted by the OECD, under the guidance of a group of mathematics experts from participating countries. Participating countries contributed stimulus material and questions, which were reviewed, tried out and refined iteratively over the three years leading up to the administration of the assessment in 2012. The development process involved provisions for several rounds of commentary from participating countries and economies, as well as small-scale piloting and a formal field trial in which samples of 15-year-olds (about 1 000 students) from participating countries and economies took part. The mathematics expert group recommended the final selection of tasks, which included material submitted by participating countries and economies. The selection was made with regard to both their technical quality, assessed on the basis of their performance in the field trial, and their cultural appropriateness and interest level for 15-year-olds, as judged by the participating countries. Another essential criterion for selecting the set of material as a whole was its fit to the framework described in Volume 1, in order to maintain the balance across various categories of context, content and process. Finally, it was carefully ensured that the set of questions covered a range of difficulty, allowing good measurement and description of the mathematics literacy of all 15-year-old students, from the least proficient to the highly able.

More than 110 print mathematics questions were used in PISA 2012, but each student in the sample only saw a fraction of the total pool because different sets of questions were given to different students. The mathematics questions selected for inclusion in PISA 2012 were organised into half-hour clusters. These, along with clusters of reading and science questions, were assembled into booklets containing four clusters each. Each participating student was then given a two-hour assessment. As mathematics was the focus of the PISA 2012 assessment, every booklet included at least one cluster of mathematics material. The clusters were rotated so that each cluster appeared in each of the four possible positions in the booklets, and each pair of clusters appeared in at least one of the 13 booklets that were used.

This design, similar to those used in previous PISA assessments, makes it possible to construct a single scale of mathematics proficiency, in which each question is associated with a particular point on the scale that indicates its difficulty, whereby each student's performance is associated with a particular point on the same scale that indicates his or her estimated proficiency. A description of the modelling technique used to construct this scale can be found in the *PISA 2012 Technical Report* (OECD, forthcoming).

The relative difficulty of tasks in a test is estimated by considering the proportion of test takers who answer each question correctly. The relative proficiency of students taking a particular test can be estimated by considering the proportion of test questions they answer correctly. A single continuous scale shows the relationship between the difficulty of questions and the proficiency of students. By constructing a scale that shows the difficulty of each question, it is possible to locate the level of mathematics literacy that the question represents. By showing the proficiency of each student on the same scale, it is possible to describe the level of mathematics literacy that the student possesses.

The location of student proficiency on this scale is set in relation to the particular group of questions used in the assessment. However, just as the sample of students taking PISA in 2012 is drawn to represent all the 15-year-olds in the participating countries and economies, so the individual questions used in the assessment are designed to represent the definition of mathematics literacy adequately. Estimates of student proficiency reflect the kinds of tasks they would be expected to perform successfully. This means that students are likely to be able to complete questions successfully at or below the difficulty level associated with their own position on the scale (but they may not always do so). Conversely, they are unlikely to be able to successfully complete questions above the difficulty level associated with their position on the scale (but they may sometimes do so).

The further a student's proficiency is located above a given question, the more likely he or she is to successfully complete the question (and other questions of similar difficulty); the further the student's proficiency is located below a given question, the lower the probability that the student will be able to successfully complete the question, and other questions of similar difficulty.

How mathematics proficiency levels are defined in PISA 2012

PISA 2012 provides an overall mathematics literacy scale, drawing on all the questions in the mathematics assessment, as well as scales for three process and four content categories. The metric for the overall mathematics scale is based on a mean for OECD countries set at 500 in PISA 2003, with a standard deviation of 100. To help interpret what students' scores mean in substantive terms, the scale is divided into levels, based on a set of statistical principles, and then descriptions are generated, based on the tasks that are located within each level, to describe the kinds of skills and knowledge needed to successfully complete those tasks.

For PISA 2012, the range of difficulty of tasks allows for the description of six levels of mathematics proficiency: Level 1 is the lowest described level, then Level 2, Level 3 and so on up to Level 6.

Students with a proficiency within the range of Level 1 are likely to be able to successfully complete Level 1 tasks (and others like them), but are unlikely to be able to complete tasks at higher levels. Level 6 reflects tasks that present the greatest challenge in terms



of mathematics skills and knowledge. Students with scores in this range are likely to be able to complete mathematics tasks located at that level successfully, as well as all the other mathematics tasks in PISA.

PISA applies a standard methodology for constructing proficiency scales. Based on a student's performance on the tasks in the test, his or her score is generated and located in a specific part of the scale, thus allowing the score to be associated with a defined proficiency level. The level at which the student's score is located is the highest level for which he or she would be expected to answer correctly most of a random selection of questions within the same level. Thus, for example, in an assessment composed of tasks spread uniformly across Level 3, students with a score located within Level 3 would be expected to complete at least 50% of the tasks successfully. Because a level covers a range of difficulty and proficiency, success rates across the band vary. Students near the bottom of the level would be likely to succeed on just over 50% of the tasks spread uniformly across the level, while students at the top of the level would be likely to succeed on well over 70% of the same tasks.

Figure I.2.21 in Volume I provides details of the nature of mathematics skills, knowledge and understanding required at each level of the mathematics scale.

Context questionnaire indices

This section explains the indices derived from the student and school context questionnaires used in PISA 2012.

Several PISA measures reflect indices that summarise responses from students, their parents or school representatives (typically principals) to a series of related questions. The questions were selected from a larger pool of questions on the basis of theoretical considerations and previous research. The *PISA 2012 Assessment and Analytical Framework* (OECD, 2013) provides an in-depth description of this conceptual framework. Structural equation modelling was used to confirm the theoretically expected behaviour of the indices and to validate their comparability across countries and economies. For this purpose, a model was estimated separately for each country and collectively for all OECD countries. For a detailed description of other PISA indices and details on the methods, see *PISA 2012 Technical Report* (OECD, forthcoming).

There are two types of indices: simple indices and scale indices.

Simple indices are the variables that are constructed through the arithmetic transformation or recoding of one or more items, in exactly the same way across assessments. Here, item responses are used to calculate meaningful variables, such as the recoding of the four-digit ISCO-08 codes into "Highest parents' socio-economic index (HISEI)" or, teacher-student ratio based on information from the school questionnaire.

Scale indices are the variables constructed through the scaling of multiple items. Unless otherwise indicated, the index was scaled using a weighted likelihood estimate (WLE) (Warm, 1989), using a one-parameter item response model (a partial credit model was used in the case of items with more than two categories). For details on how each scale index was constructed see the *PISA 2012 Technical Report* (OECD, forthcoming). In general, the scaling was done in three stages:

- The item parameters were estimated from equal-sized subsamples of students from all participating countries and economies.
- The estimates were computed for all students and all schools by anchoring the item parameters obtained in the preceding step.
- The indices were then standardised so that the mean of the index value for the OECD student population was zero and the standard deviation was one (countries being given equal weight in the standardisation process).

Sequential codes were assigned to the different response categories of the questions in the sequence in which the latter appeared in the student, school or parent questionnaires. Where indicated in this section, these codes were inverted for the purpose of constructing indices or scales. Negative values for an index do not necessarily imply that students responded negatively to the underlying questions. A negative value merely indicates that the respondents answered less positively than all respondents did on average across OECD countries. Likewise, a positive value on an index indicates that the respondents answered more favourably, or more positively, than respondents did, on average, across OECD countries. Terms enclosed in brackets < > in the following descriptions were replaced in the national versions of the student, school and parent questionnaires by the appropriate national equivalent. For example, the term <qualification at ISCED level 5A> was translated in the United States into "Bachelor's degree, post-graduate certificate program, Master's degree program or first professional degree program". Similarly the term <classes in the language of assessment> in Luxembourg was translated into "German classes" or "French classes" depending on whether students received the German or French version of the assessment instruments.

In addition to simple and scaled indices described in this annex, there are a number of variables from the questionnaires that correspond to single items not used to construct indices. These non-recoded variables have prefix of "ST" for the questionnaire items in the student questionnaire, "SC" for the items in the school questionnaire, and "PA" for the items in the parent questionnaire. All the context questionnaires as well as the PISA international database, including all variables, are available through www.pisa.oecd.org.

Scaling of questionnaire indices for trend analyses

In PISA, to gather information about students' and schools' characteristics, both students and schools complete a background questionnaire. In PISA 2003 and PISA 2012 several questions were kept untouched, enabling the comparison of responses to these



questions over time. In this report, only questions that maintained an exact wording are used for trends analyses. Questions with subtle word changes or questions with major word changes were not compared across time because it is impossible to discern whether observed changes in the response are due to changes in the construct they are measuring or to changes in the way the construct is being measured.

Also, in PISA, as described in this Annex, questionnaire items are used to construct indices. Whenever the questions used in the construction of indices remains intact in PISA 2003 and PISA 2012, the corresponding indices are compared. Two types of indices are used in PISA: simple indices and scale indices.

Simple indices recode a set of responses to questionnaire items. For trends analyses, the values observed in PISA 2003 are compared directly to PISA 2012, just as simple responses to questionnaire items are. This is the case of indices like student-teacher ratio and ability grouping in mathematics.

Scale indices, on the other hand, imply WLE estimates which require rescaling in order to be comparable across PISA cycles. Scale indices, like the *PISA index of economic, social and cultural status*, the *index of sense of belonging*, the *index of attitudes towards school*, the *index of intrinsic motivation to learn mathematics*, the *index of instrumental motivation to learn mathematics*, the *index of mathematics self-efficacy*, the *index of mathematics self-concept*, the *index of anxiety towards mathematics*, the *index of teacher shortage*, the *index of quality of physical infrastructure*, the *index of quality of educational resources*, the *index of disciplinary climate*, the *index of teacher-student relations*, the *index of teacher morale*, the *index of student-related factors affecting school climate* and the *index of teacher-related factors affecting school climate*, were scaled, in PISA 2012 to have an OECD average of 0 and a standard deviation of 1, on average, across OECD countries. These same scales were scaled, in PISA 2003, to have an OECD average of 0 and a standard deviation of 1. Because they are on different scales, values reported in *Learning for Tomorrow's World: First Results from PISA 2003* (OECD, 2004) cannot be compared with those reported in this volume. To make these scale indices comparable, values for 2003 have been rescaled to the 2012 scale, using the PISA 2012 parameter estimates.

These re-scaled indices are available at www.pisa.oecd.org. They can be merged to the corresponding PISA 2003 dataset using the country names, school and student-level identifiers. The rescaled *PISA index of economic, social and cultural status* is also available to be merged with the PISA 2000, PISA 2006 and PISA 2009 dataset.

Student-level simple indices

Age

The variable AGE is calculated as the difference between the middle month and the year in which students were assessed and their month and year of birth, expressed in years and months.

Study programme

In PISA 2012, study programmes available to 15-year-old students in each country were collected both through the student tracking form and the student questionnaire (ST02). All study programmes were classified using ISCED (OECD, 1999). In the PISA international database, all national programmes are indicated in a variable (PROGN) where the first six digits refer to the national centre code and the last two digits to the national study programme code.

The following internationally comparable indices were derived from the data on study programmes:

- Programme level (ISCEDL) indicates whether students are (1) primary education level (ISCED 1); (2) lower-secondary education level; or (3) upper secondary education level.
- Programme designation (ISCEDD) indicates the designation of the study programme: (1) "A" (general programmes designed to give access to the next programme level); (2) "B" (programmes designed to give access to vocational studies at the next programme level); (3) "C" (programmes designed to give direct access to the labour market); or (4) "M" (modular programmes that combine any or all of these characteristics).
- Programme orientation (ISCEDO) indicates whether the programme's curricular content is (1) general; (2) pre-vocational; (3) vocational; or (4) modular programmes that combine any or all of these characteristics.

Occupational status of parents

Occupational data for both a student's father and a student's mother were obtained by asking open-ended questions in the student questionnaire (ST12, ST16). The responses were coded to four-digit ISCO codes (ILO, 1990) and then mapped to the SEI index of Ganzeboom et al. (1992). Higher scores of SEI indicate higher levels of occupational status. The following three indices are obtained:

- Mother's occupational status (OCOD1).
- Father's occupational status (OCOD2).
- The highest occupational level of parents (HISEI) corresponds to the higher SEI score of either parent or to the only available parent's SEI score.




[Part 1/1]
Table A1.1 Levels of parental education converted into years of schooling

	Completed ISCED level 1 (primary education)	Completed ISCED level 2 (lower secondary education)	Completed ISCED levels 3B or 3C (upper secondary education providing direct access to the labour market or to ISCED 5B programmes)	Completed ISCED level 3A (upper secondary education providing access to ISCED 5A and 5B programmes) and/ or ISCED level 4 (non- tertiary post-secondary)	Completed ISCED level 5A (university level tertiary education) or ISCED level 6 (advanced research programmes)	Completed ISCED level 5B (non-university tertiary education)
OECD						
Australia	6.0	10.0	11.0	12.0	15.0	14.0
Austria	4.0	9.0	12.0	12.5	17.0	15.0
Belgium ¹	6.0	9.0	12.0	12.0	17.0	15.0
Canada	6.0	9.0	12.0	12.0	17.0	15.0
Chile	6.0	8.0	12.0	12.0	17.0	16.0
Czech Republic	5.0	9.0	11.0	13.0	16.0	16.0
Denmark	7.0	10.0	13.0	13.0	18.0	16.0
Estonia	6.0	9.0	12.0	12.0	16.0	15.0
Finland	6.0	9.0	12.0	12.0	16.5	14.5
France	5.0	9.0	12.0	12.0	15.0	14.0
Germany	4.0	10.0	13.0	13.0	18.0	15.0
Greece	6.0	9.0	11.5	12.0	17.0	15.0
Hungary	4.0	8.0	10.5	12.0	16.5	13.5
Iceland	7.0	10.0	13.0	14.0	18.0	16.0
Ireland	6.0	9.0	12.0	12.0	16.0	14.0
Israel	6.0	9.0	12.0	12.0	15.0	15.0
Italy	5.0	8.0	12.0	13.0	17.0	16.0
Japan	6.0	9.0	12.0	12.0	16.0	14.0
Korea	6.0	9.0	12.0	12.0	16.0	14.0
Luxembourg	6.0	9.0	12.0	13.0	17.0	16.0
Mexico	6.0	9.0	12.0	12.0	16.0	14.0
Netherlands	6.0	10.0	13.0	12.0	16.0	15.0
New Zealand	5.5	10.0	11.0	12.0	15.0	14.0
Norway	6.0	9.0	12.0	12.0	16.0	14.0
Poland	a	8.0	11.0	12.0	16.0	15.0
Portugal	6.0	9.0	12.0	12.0	17.0	15.0
Slovak Republic ²	4.0	9.0	12.0	13.0	18.0	16.0
Slovenia	4.0	8.0	11.0	12.0	16.0	15.0
Spain	5.0	8.0	10.0	12.0	16.5	13.0
Sweden	6.0	9.0	11.5	12.0	16.0	14.0
Switzerland	6.0	9.0	12.5	12.5	17.5	14.5
Turkey	5.0	8.0	11.0	11.0	15.0	13.0
United Kingdom (exclud. Scotland)	6.0	9.0	12.0	13.0	16.0	15.0
United Kingdom (Scotland)	7.0	9.0	11.0	13.0	17.0	15.0
United States	6.0	9.0	a	12.0	16.0	14.0
Partners						
Albania	6.0	9.0	12.0	12.0	16.0	16.0
Argentina	6.0	10.0	12.0	12.0	17.0	14.5
Azerbaijan	4.0	9.0	11.0	11.0	17.0	14.0
Brazil	4.0	8.0	11.0	11.0	16.0	14.5
Bulgaria	4.0	8.0	10.0	12.0	17.5	15.0
Colombia	5.0	9.0	11.0	11.0	15.5	14.0
Costa Rica	6.0	9.0	11.0	12.0	14.0	16.0
Croatia	4.0	8.0	11.0	12.0	17.0	15.0
Hong Kong-China	6.0	9.0	11.0	13.0	16.0	14.0
Indonesia	6.0	9.0	12.0	12.0	15.0	14.0
Jordan	6.0	10.0	12.0	12.0	16.0	14.5
Kazakhstan	4.0	9.0	11.5	12.5	15.0	14.0
Latvia	4.0	8.0	11.0	11.0	16.0	14.0
Liechtenstein	5.0	9.0	11.0	13.0	17.0	14.0
Lithuania	3.0	8.0	11.0	11.0	16.0	15.0
Macao-China	6.0	9.0	11.0	12.0	16.0	15.0
Malaysia	6.0	9.0	11.0	13.0	15.0	16.0
Montenegro	4.0	8.0	11.0	12.0	16.0	15.0
Peru	6.0	9.0	11.0	11.0	17.0	14.0
Qatar	6.0	9.0	12.0	12.0	16.0	15.0
Romania	4.0	8.0	11.5	12.5	16.0	14.0
Russian Federation	4.0	9.0	11.5	12.0	15.0	a
Serbia	4.0	8.0	11.0	12.0	17.0	14.5
Shanghai-China	6.0	9.0	12.0	12.0	16.0	15.0
Singapore	6.0	8.0	10.0	11.0	16.0	13.0
Chinese Taipei	6.0	9.0	12.0	12.0	16.0	14.0
Thailand	6.0	9.0	12.0	12.0	16.0	14.0
Tunisia	6.0	9.0	12.0	13.0	17.0	16.0
United Arab Emirates	5.0	9.0	12.0	12.0	16.0	15.0
Uruguay	6.0	9.0	12.0	12.0	17.0	15.0
Viet Nam	5.0	9.0	12.0	12.0	17.0	a

1. In Belgium the distinction between universities and other tertiary schools doesn't match the distinction between ISCED 5A and ISCED 5B.

2. In the Slovak Republic, university education (ISCED 5A) usually lasts five years and doctoral studies (ISCED 6) lasts three more years. Therefore, university graduates will have completed 18 years of study and graduates of doctoral programmes will have completed 21 years of study.

Source: OECD, PISA 2012 Database.

StatLink  <http://dx.doi.org/10.1787/888932937073>



Some of the analyses distinguish between four different categories of occupations by the major groups identified by the ISCO coding of the highest parental occupation: Elementary (ISCO 9), semi-skilled blue-collar (ISCO 6, 7 and 8), semi-skilled white-collar (ISCO 4 and 5), skilled (ISCO 1, 2 and 3). This classification follows the same methodology used in other OECD publications such as *Education at a Glance* (2013b) and the *OECD Skills Outlook* (2013c).¹

Educational level of parents

The educational level of parents is classified using ISCED (OECD, 1999) based on students' responses in the student questionnaire (ST13, ST14, ST17 and ST18).

As in PISA 2000, 2003, 2006 and 2009, indices were constructed by selecting the highest level for each parent and then assigning them to the following categories: (0) None, (1) ISCED 1 (primary education), (2) ISCED 2 (lower secondary), (3) ISCED Level 3B or 3C (vocational/pre-vocational upper secondary), (4) ISCED 3A (upper secondary) and/or ISCED 4 (non-tertiary post-secondary), (5) ISCED 5B (vocational tertiary), (6) ISCED 5A, 6 (theoretically oriented tertiary and post-graduate). The following three indices with these categories are developed:

- Mother's educational level (MISCED).
- Father's educational level (FISCED).
- Highest educational level of parents (HISCED) corresponds to the higher ISCED level of either parent.

Highest educational level of parents was also converted into the number of years of schooling (PARED). For the conversion of level of education into years of schooling, see Table A1.1.

Immigration and language background

Information on the country of birth of students and their parents is collected in a similar manner as in PISA 2000, PISA 2003, PISA 2006 and PISA 2009 by using nationally specific ISO coded variables. The ISO codes of the country of birth for students and their parents are available in the PISA international database (COBN_S, COBN_M, and COBN_F).

The index on immigrant background (IMMIG) has the following categories: (1) non-immigrant students (those students born in the country of assessment, or those with at least one parent born in that country; students who were born abroad with at least one parent born in the country of assessment are also classified as non-immigrant students), (2) second-generation students (those born in the country of assessment but whose parents were born in another country) and (3) first-generation students (those born outside the country of assessment and whose parents were also born in another country). Students with missing responses for either the student or for both parents, or for all three questions have been given missing values for this variable.

Students indicate the language they usually speak at home. The data are captured in nationally-specific language codes, which were recoded into variable LANGN with the following two values: (1) language at home is the same as the language of assessment, and (2) language at home is a different language than the language of assessment.

Relative grade

Data on the student's grade are obtained both from the student questionnaire (ST01) and from the student tracking form. As with all variables that are on both the tracking form and the questionnaire, inconsistencies between the two sources are reviewed and resolved during data-cleaning. In order to capture between-country variation, the relative grade index (GRADE) indicates whether students are at the modal grade in a country (value of 0), or whether they are below or above the modal grade level (+ x grades, - x grades).

The relationship between the grade and student performance was estimated through a multilevel model accounting for the following background variables: *i*) the *PISA index of economic, social and cultural status*; *ii*) the *PISA index of economic, social and cultural status squared*; *iii*) the school mean of the *PISA index of economic, social and cultural status*; *iv*) an indicator as to whether students were foreign-born first-generation students; *v*) the percentage of first-generation students in the school; and *vi*) students' gender.

Table A1.2 presents the results of the multilevel model. Column 1 in Table A1.2 estimates the score-point difference that is associated with one grade level (or school year). This difference can be estimated for the 32 OECD countries in which a sizeable number of 15-year-olds in the PISA samples were enrolled in at least two different grades. Since 15-year-olds cannot be assumed to be distributed at random across the grade levels, adjustments had to be made for the above-mentioned contextual factors that may relate to the assignment of students to the different grade levels. These adjustments are documented in columns 2 to 7 of the table. While it is possible to estimate the typical performance difference among students in two adjacent grades net of the effects of selection and contextual factors, this difference cannot automatically be equated with the progress that students have made over the last school year but should be interpreted as a lower boundary of the progress achieved. This is not only because different students were assessed but also because the content of the PISA assessment was not expressly designed to match what students had learned in the preceding school year but more broadly to assess the cumulative outcome of learning in school up to age 15. For example, if the curriculum of the grades in which 15-year-olds are enrolled mainly includes material other than that assessed by PISA (which, in turn, may have been included in earlier school years) then the observed performance difference will underestimate student progress.



Learning time

Learning time in test language (LMINS) was computed by multiplying students' responses on the number of minutes on average in the test language class by number of test language class periods per week (ST69 and ST70). Comparable indices were computed for mathematics (MMINS) and science (SMINS).

Engagement with and at school

Skiping classes or days of school

Student responses over whether, in the two weeks before the PISA test, they skipped classes (ST09) or days of school (ST115) at least once were used to derive an indicator of student truancy which takes value 0 if students reported not skipping any class and not skipping any day of school in the two weeks before the PISA test and value 1 if students reported skipping classes or days of school at least once in the same period.

Sense of belonging

The *index of sense of belonging* (BELONG) was constructed using student responses (ST87) over the extent they strongly agreed, agreed, disagreed or strongly disagreed to the following statements: I feel like an outsider (or left out of things) at school; I make friends easily at school; I feel like I belong at school; I feel awkward or out of place in my school; other students seem to like me; I feel lonely at school; I feel happy at school; things are ideal in my school; I am satisfied with my school.

For trends analyses, the PISA 2003 values of the index of sense of belonging were rescaled to be comparable to those in PISA 2012. As a result, values for the index of sense of belonging for PISA 2003 reported in this volume may differ from those reported in *Learning for Tomorrow's World: First Results from PISA 2003* (OECD, 2004). Three of the questions included to compute the index of sense of belonging in PISA 2012 ("I feel happy at school," "things are ideal in my school," and "I am satisfied with my school") were not included in the PISA 2003 questionnaire. Estimation of the PISA 2003 index treats these questions as missing and, under the assumption that the relationship between the items remains unchanged with the inclusion of the new questions, the PISA 2003 and PISA 2012 values on the index of sense of belonging are comparable after the rescaling.

Attitudes towards school (learning outcomes)

The *index of attitudes towards school (learning outcomes)* (ATSCHL) was constructed using student responses (ST88) over the extent they strongly agreed, agreed, disagreed or strongly disagreed to the following statements when asked about what they have learned in school: School has done little to prepare me for adult life when I leave school; school has been a waste of time; school has helped give me confidence to make decisions; school has taught me things which could be useful in a job.

For trends analyses, the PISA 2003 values of the index of attitudes towards school were rescaled to be comparable to those in PISA 2012. As a result, values for the index of attitudes towards school for PISA 2003 reported in this volume may differ from those reported in *Learning for Tomorrow's World: First Results from PISA 2003* (OECD, 2004).

Attitudes towards school (learning activities)

The *index of attitudes towards school (learning activities)* (ATLNACT) was constructed using student responses (ST89) over the extent they strongly agreed, agreed, disagreed or strongly disagreed to the following statements when asked to think about their school: Trying hard at school will help me get a good job; trying hard at school will help me get into a good <college>; I enjoy receiving good <grades>; trying hard at school is important.

Student drive and motivation

Perseverance

The *index of perseverance* (PERSEV) was constructed using student responses (ST93) over whether they report that the following statements describe them very much, mostly, somewhat, not much, not at all: When confronted with a problem, I give up easily; I put off difficult problems; I remain interested in the tasks that I start; I continue working on tasks until everything is perfect; when confronted with a problem, I do more than what is expected of me.

Openness to problem solving

The *index of openness to problem solving* (OPENPS) was constructed using student responses (ST94) over whether they report that the following statements describe them very much, mostly, somewhat, not much, not at all: I can handle a lot of information; I am quick to understand things; I seek explanations of things; I can easily link facts together; I like to solve complex problems.

Perceived self-responsibility for failing in mathematics

The *index of perceived self-responsibility for failing in mathematics* (FAILMAT) was constructed using student responses when examining the following scenario defined in (ST44): "suppose that you are a student in the following situation: each week, your mathematics teacher gives a short quiz. Recently you have done badly on these quizzes. Today you are trying to figure out why. Are you very likely, likely, slightly likely or not at all likely to have the following thoughts or feelings in this situation? I'm not very good at solving



mathematics problems; my teacher did not explain the concepts well this week; this week I made bad guesses on the quiz; sometimes the course material is too hard; the teacher did not get students interested in the material; sometimes I am just unlucky.

Intrinsic motivation to learn mathematics

The *index of intrinsic motivation to learn mathematics* (INTMAT) was constructed using student responses over the extent they strongly agreed, agreed, disagreed or strongly disagreed to the statements asked in question (ST29), when asked to think about their views on mathematics: I enjoy reading about mathematics; I look forward to my mathematics; I do mathematics because I enjoy it; I am interested in the things I learn in mathematics.

For trends analyses, the PISA 2003 values of the index of intrinsic motivation to learn mathematics were rescaled to be comparable to those in PISA 2012. As a result, values for the index of intrinsic motivation to learn mathematics for PISA 2003 reported in this volume may differ from those reported in *Learning for Tomorrow's World: First Results from PISA 2003* (OECD, 2004). In PISA 2003 the index of intrinsic motivation to learn mathematics was named the index of interest and enjoyment in mathematics. Given that both are based on the same questionnaire items, they are comparable over time.

Instrumental motivation to learn mathematics

The *index of instrumental motivation to learn mathematics* (INSTMOT) was constructed using student responses over the extent they strongly agreed, agreed, disagreed or strongly disagreed to a series of statements in question (ST29) when asked to think about their views on mathematics: Making an effort in mathematics is worth because it will help me in the work that I want to do later on; learning mathematics is worthwhile for me because it will improve my career <prospects, chances>; Mathematics is an important subject for me because I need it for what I want to study later on; I will learn many things in mathematics that will help me get a job.

For trends analyses, the PISA 2003 values of the index of instrumental motivation to learn mathematics were rescaled to be comparable to those in PISA 2012. As a result, values for the index of instrumental motivation to learn mathematics for PISA 2003 reported in this volume may differ from those reported in *Learning for Tomorrow's World: First Results from PISA 2003* (OECD, 2004).

Mathematics self-beliefs

Mathematics self-efficacy

The *index of mathematics self-efficacy* (MATHEFF) was constructed using student responses over the extent they reported feeling very confident, confident, not very confident, not at confident about having to do a number of tasks. The question (ST37) asked about the following mathematics tasks: Using a <train timetable> to work out how long it would take to get from one place to another; calculating how much cheaper a TV would be after a 30% discount; calculating how many square metres of tiles you need to cover a floor; understanding graphs presented in newspapers; solving an equation like $3x+5=17$; finding the actual distance between two places on a map with a 1:10,000 scale; solving an equation like $2(x+3)=(x+3)(x-3)$; calculating the petrol consumption rate of a car. Making an effort in mathematics is worth because it will help me in the work that I want to do later on; learning mathematics is worthwhile for me because it will improve my career <prospects, chances>; Mathematics is an important subject for me because I need it for what I want to study later on; I will learn many things in mathematics that will help me get a job.

For trends analyses, the PISA 2003 values of the index of mathematics self-efficacy were rescaled to be comparable to those in PISA 2012. As a result, values for the index of mathematics self-efficacy for PISA 2003 reported in this volume may differ from those reported in *Learning for Tomorrow's World: First Results from PISA 2003* (OECD, 2004).

Mathematics self-concept

The *index of mathematics self-concept* (SCMAT) was constructed using student responses to question (ST42) over the extent they strongly agreed, agreed, disagreed or strongly disagreed with the following statements when asked to think about studying mathematics: I am just not good at mathematics; I get good <grades> in mathematics; I learn mathematics quickly; I have always believed that mathematics is one of my best subjects; in my mathematics class, I understand even the most difficult work.

For trends analyses, the PISA 2003 values of the index of mathematics self-concept were rescaled to be comparable to those in PISA 2012. As a result, values for the index of mathematics self-concept for PISA 2003 reported in this volume may differ from those reported in *Learning for Tomorrow's World: First Results from PISA 2003* (OECD, 2004).

Mathematics anxiety

The *index of mathematics anxiety* (ANXMAT) was constructed using student responses to question (ST42) over the extent they strongly agreed, agreed, disagreed or strongly disagreed with the following statements when asked to think about studying mathematics: I often worry that it will be difficult for me in mathematics classes; I get very tense when I have to do mathematics homework; I get very nervous doing mathematics problems; I feel helpless when doing a mathematics problem; I worry that I will get poor <grades> in mathematics.

Table A1.3 Student questionnaire rotation design

Form A	Common Question Set (all forms)	Question Set 1 – Mathematics Attitudes/ Problem Solving	Question Set 3 – Opportunity to Learn/ Learning Strategies
Form B	Common Question Set (all forms)	Question Set 2 – School Climate/Attitudes towards School/Anxiety	Question Set 1 – Mathematics Attitudes/ Problem Solving
Form C	Common Question Set (all forms)	Question Set 3 – Opportunity to Learn/ Learning Strategies	Question Set 2 – School Climate/Attitudes towards School/Anxiety

Note: For details regarding the questions in each question set, please refer to the *PISA 2012 Technical Report* (OECD, forthcoming).

For trends analyses, the PISA 2003 values of the index of anxiety towards mathematics were rescaled to be comparable to those in PISA 2012. As a result, values for the index of anxiety towards mathematics for PISA 2003 reported in this volume may differ from those reported in *Learning for Tomorrow's World: First Results from PISA 2003* (OECD, 2004).

Dispositions towards mathematics

Mathematics intentions

The index of mathematics intentions (MATINTFC) was constructed asking students (ST48) to choose, for each pair of the following statements, the item that best described them: I intend to take additional mathematics courses after school finishes vs. I intend to take additional <test language> courses after school finishes; I plan on majoring in a subject in <college> that requires mathematics skills vs. I plan on majoring in a subject in <college> that requires science skills; I am willing to study harder in my mathematics classes than is required vs. I am willing to study harder in my <test language> classes than is required; I plan on <taking> as many mathematics classes as I can during my education vs. I plan on <taking> as many science classes as I can during my education; I am planning on pursuing a career that involves a lot of mathematics vs. I am planning on pursuing a career that involves a lot of science.

Subjective norms in mathematics

The index of subjective norms in mathematics (SUBNORM) was constructed using student responses (ST35) over whether, thinking about how people important to them view mathematics, they strongly agreed, agreed, disagreed or strongly disagreed to the following statements: Most of my friends do well in mathematics; most of my friends work hard at mathematics; my friends enjoy taking mathematics tests; my parents believe it's important for me to study mathematics; my parents believe that mathematics is important for my career; my parents like mathematics

Mathematics behaviours

The index of mathematics behaviours (MATBEH) was constructed using student responses (ST49) over how often (always or almost always, often, sometimes, never, rarely) they do the following things at school and outside of school: I talk about mathematics problems with my friends; I help my friends with mathematics; I do mathematics as an <extracurricular> activity; I take part in mathematics competitions; I do mathematics more than 2 hours a day outside of school; I play chess; I program computers; I participate in a mathematics club.

Pre-primary attendance

Students were asked (ST05) whether they had attended pre-primary education (<ISCED 0> which was the adapted by each national centre) and if they did, if they attended "for one year or less" or "for more than one year."

Mother/father Current Job Status

After answering questions about parental occupation and education, students were asked (ST15 and ST19) "What is your mother/father currently doing?". They could then choose between four options: *i*) "Working full-time" <for pay>, *ii*) Working part-time <for pay>, *iii*) Not working, but looking for a job and, *iv*) Other (e.g. home duties, retired).

Teacher-Directed instruction

The index of teacher-directed instruction (TCHBEHTD) was constructed using students' reports (ST79) on the frequency (every lesson, most lessons, some lessons, never or hardly ever) with which, in mathematics lessons, the teacher sets clear goals for student learning; the teacher asks students to present their thinking or reasoning at some length; the teacher asks questions to check whether students understood what was taught; and the teacher tells students what they have to learn.

Teachers' student orientation

The index of teachers' student orientation (TCHBEHSO) was constructed using students' reports (ST79) on the frequency (every lesson, most lessons, some lessons, never or hardly ever) with which, in mathematics lessons, the teacher gives students different work to classmates who have difficulties learning and/or to those who can advance faster; the teacher assigns projects that require at least one week to complete; the teacher has students work in small groups to come up with a joint solution to a problem or task; and the teacher asks students to help plan classroom activities or topics.



Teachers' use of formative assessment

The *index of teachers' use of formative assessment* (TCHBEHFA) was constructed using students' reports (ST79) on the frequency (every lesson, most lessons, some lessons, never or hardly ever) with which, in mathematics lessons, the teacher tells students how well they are doing in mathematics class; the teacher gives students feedback on their strengths and weaknesses in mathematics; and the teacher tells students what they need to do to become better in mathematics.

Cognitive activation

The *index of teacher's use of cognitive activation strategies* (COGACT) was constructed using student responses (ST80) over how often (always or almost always, often, sometimes, never or rarely) a series of situations happened with the mathematics teacher that taught them their last mathematics class: the teacher asks questions that make students reflect on the problem; the teacher gives problems that require students to think for an extended time; the teacher asks students to decide, on their own, procedures for solving complex problems; the teacher presents problems in different contexts so that students know whether they have understood the concepts; the teacher helps students to learn from mistakes they have made; the teacher asks students to explain how they solved a problem; the teacher presents problems that require students to apply what they have learned in new contexts; and the teacher gives problems that can be solved in different ways. Students were asked to report whether these behaviours and situations occur always or almost always, often, sometimes or never or rarely.

Experience with applied mathematics tasks

The *index of student experience with applied mathematics tasks* (EXAPPLM) was constructed using student responses (ST61) on whether they have frequently, sometimes, rarely or never encountered the following types of mathematics tasks during their time at school: working out from a <train timetable> how long it would take to get from one place to another; calculating how much more expensive a computer would be after adding tax; calculating how many square metres of tiles you need to cover a floor; understanding scientific tables presented in an article; finding the actual distance between two places on a map with a 1:10,000 scale; calculating the power consumption of an electronic appliance per week.

Experience with pure mathematics tasks

The *index of student experience with pure mathematics tasks* (EXPUREM) was constructed using student responses (ST61) on whether they have frequently, sometimes, rarely or never encountered the following types of mathematics tasks during their time at school: solving an equation like $6x^2+5=29$; solving an equation like $2(x+3)=(x+3)(x-3)$; solving an equation like $3x+5=17$.

For trends analyses, the PISA 2003 values of the index of student-teacher relations were rescaled to be comparable to those in PISA 2012. As a result, values for the index of student-teacher relations for PISA 2003 reported in this volume may differ from those reported in *Learning for Tomorrow's World: First Results from PISA 2003* (OECD, 2004).

Parental occupation in STEM fields

Student responses over their parents' occupation were used to develop the index of parental occupation in STEM fields. The index takes value 1 if at least one parent works in ISCO-08 occupations: 2100 to 2166 (Science and engineering professionals, Physical and earth science professionals, Physicists and astronomers, Meteorologists, Chemists, Geologists and geophysicists, Mathematicians, actuaries and statisticians, Life science professionals, Biologists, botanists, zoologists and related professionals, Farming, forestry and fisheries advisers, Environmental protection professionals, Engineering professionals (excluding electrotechnology), Industrial and production engineers, Civil engineers, Environmental engineers, Mechanical engineers, Chemical engineers, Mining engineers, metallurgists and related professionals, Engineering professionals not elsewhere classified, Electrotechnology engineers, Electrical engineers, Electronics engineers, Telecommunications engineers, Architects, planners, surveyors and designers, Building architects, Landscape architects, Product and garment designers, Town and traffic planners, Cartographers and surveyors, Graphic and multimedia designers); 2510 to 2529 (Software and applications developers and analysts, Systems analysts, Software developers, Web and multimedia developers, Applications programmers, Software and applications developers and analysts not elsewhere classified, Database and network professionals, Database designers and administrators, Systems administrators, Computer network professionals, Database and network professionals not elsewhere classified); 3100 to 3155 (Science and engineering associate professionals, Physical and engineering science technicians, Chemical and physical science technicians, Civil engineering technicians, Electrical engineering technicians, Electronics engineering technicians, Mechanical engineering technicians, Chemical engineering technicians, Mining and metallurgical technicians, Draughtspersons, Physical and engineering science technicians not elsewhere classified, Mining, manufacturing and construction supervisors, Mining supervisors, Manufacturing supervisors, Construction supervisors, Process control technicians, Power production plant operators, Incinerator and water treatment plant operators, Chemical processing plant controllers, Petroleum and natural gas refining plant operators, Metal production process controllers, Process control technicians not elsewhere classified, Life science technicians and related associate professionals, Life science technicians (excluding medical), Agricultural technicians, Forestry technicians, Ship and aircraft controllers and technicians, Ships' engineers, Ships' deck officers and pilots, Aircraft pilots and related associate professionals, Air traffic controllers, Air traffic safety electronics technicians); 2631 (Economists); 3314 (Statistical, mathematical and related associate professionals); and 2413 (Financial analysts).



Disciplinary climate

The *index of disciplinary climate* (DISCLIMA) was derived from students' reports on how often the followings happened in their lessons of the language of instruction (ST81): *i*) students don't listen to what the teacher says; *ii*) there is noise and disorder; *iii*) the teacher has to wait a long time for the students to <quieten down>; *iv*) students cannot work well; and *v*) students don't start working for a long time after the lesson begins. In this index higher values indicate a better disciplinary climate.

For trends analyses, the PISA 2003 values of the index of disciplinary climate were rescaled to be comparable to those in PISA 2012. As a result, values for the index of disciplinary climate for PISA 2003 reported in this volume may differ from those reported in *Learning for Tomorrow's World: First Results from PISA 2003* (OECD, 2004).

Teacher-student relations

The *index of teacher-student relations* (STUDREL) was derived from students' level of agreement with the following statements. The question asked (ST86) stated "Thinking about the teachers at your school: to what extent do you agree with the following statements": *i*) Students get along well with most of my teachers; *ii*) Most teachers are interested in students' well-being; *iii*) Most of my teachers really listen to what I have to say; *iv*) if I need extra help, I will receive it from my teachers; and *v*) Most of my teachers treat me fairly. Higher values on this index indicate positive teacher-student relations.

For trends analyses, the PISA 2003 values of the index of student-teacher relations were rescaled to be comparable to those in PISA 2012. As a result, values for the index of student-teacher relations for PISA 2003 reported in this volume may differ from those reported in *Learning for Tomorrow's World: First Results from PISA 2003* (OECD, 2004).

Economic, social and cultural status

The *PISA index of economic, social and cultural status* (ESCS) was derived from the following three indices: *highest occupational status of parents* (HISEI), *highest educational level of parents* in years of education according to ISCED (PARED), and *home possessions* (HOMEPOS). The *index of home possessions* (HOMEPOS) comprises all items on the indices of WEALTH, CULTPOSS and HEDRES, as well as books in the home recoded into a four-level categorical variable (0-10 books, 11-25 or 26-100 books, 101-200 or 201-500 books, more than 500 books).

The *PISA index of economic, social and cultural status* (ESCS) was derived from a principal component analysis of standardised variables (each variable has an OECD mean of zero and a standard deviation of one), taking the factor scores for the first principal component as measures of the *PISA index of economic, social and cultural status*.

Principal component analysis was also performed for each participating country or economy to determine to what extent the components of the index operate in similar ways across countries or economy. The analysis revealed that patterns of factor loading were very similar across countries, with all three components contributing to a similar extent to the index (for details on reliability and factor loadings, see the *PISA 2012 Technical Report* (OECD, forthcoming).

The imputation of components for students with missing data on one component was done on the basis of a regression on the other two variables, with an additional random error component. The final values on the *PISA index of economic, social and cultural status* (ESCS) for 2012 have an OECD mean of 0 and a standard deviation of one.

ESCS was computed for all students in the five cycles, and ESCS indices for trends analyses were obtained by applying the parameters used to derive standardised values in 2012 to the ESCS components for previous cycles. These values will therefore not be directly comparable to ESCS values in the databases for previous cycles, though the differences are not large for the 2006 and 2009 cycles. ESCS values in earlier cycles were computed using different algorithms, so for 2000 and 2003 the differences are larger.

Changes to the computation of socio-economic status for PISA 2012

While the computation of socio-economic status followed what had been done in previous cycles, PISA 2012 undertook an important upgrade with respect to the coding of parental occupation. Prior to PISA 2012, the 1988 International Standard Classification of Occupations (ISCO-88) was used for the coding of parental occupation. By 2012, however, ISCO-88 was almost 25 years old and it was no longer tenable to maintain its use as an occupational coding scheme.² It was therefore decided to use its replacement, ISCO-08, for occupational coding in PISA 2012.

The change from ISCO-88 to ISCO-08 required an update of the International Socio-Economic Index (ISEI) of occupation codes. PISA 2012 therefore used a modified quantification scheme for ISCO-08 (referred to as ISEI-08), as developed by Harry Ganzeboom (2010). ISEI-08 was constructed using a database of 198 500 men and women with valid education, occupation and (personal) incomes derived from the combined 2002-07 datasets of the International Social Survey Programme (ISSP) (Ganzeboom, 2010). The methodology used for this purpose was similar to the one employed in the construction of ISEI for ISCO-68 and ISCO-88 described in different publications (Ganzeboom, de Graff and Treiman, 1992; Ganzeboom and Treiman, 1996; Ganzeboom and Treiman, 2003).³

The main differences with regard to the previous ISEI construction are the following:



- A new database was used which is more recent, larger and cross-nationally more diverse than the one used earlier.
- The new ISEI was constructed using data for women and men, while previously only men were used to estimate the scale. The data on income were corrected for hours worked to adjust the different prevalence of part-time work between men and women in many countries.

A range of validation activities accompanied the transition from ISCO-88/ISEI-88 to ISCO-08/ISEI-08, including a comparison of *i*) the distributions of ISEI-88 with ISEI-08 in terms of range, mean and standard deviations for both mothers' and fathers' occupations and *ii*) correlations between the two ISEI indicators and performance, again separately undertaken for mothers' and fathers' occupations.

For this cycle, in order to obtain trends for all cycles from 2000 to 2012, the computation of the indices WEALTH, HEDRES, CULTPOSS and HOMEPOS was based on data from all cycles from 2000 to 2012. HOMEPOS is of particular importance as it is used in the computation of ESCS. These were then standardised on 2012 so that the OECD mean is 0 and the standard deviation is 1. This means that the indices calculated on the previous cycle will be on the 2012 scale and thus not directly comparable to the indices in the database for the previously released cycles. To estimate item parameters for scaling, a calibration sample from all cycles was used, consisting of 500 students from all countries in the previous cycles, and 750 from 2012, as any particular student questionnaire item only occurs in two-thirds of the questionnaires in 2012.

The items used in the computation of the indices has changed to some extent from cycle to cycle, though cycles they have remained much the same from 2006 to 2012. The earlier cycles were are in general missing a few items that are present in the later cycles, but it was felt leaving out items only present in the later cycles would give too much weight to the earlier cycles. So a superset of all items (except country specific items) in the five cycles was used, and international item parameters were derived from this set.

The second step was to estimate WLEs for the indices, anchoring parameters on the international item set while estimating the country specific item parameters. This is the same procedure used in previous cycles.

Family wealth

The *index of family wealth* (WEALTH) is based on students' responses on whether they had the following at home: a room of their own, a link to the Internet, a dishwasher (treated as a country-specific item), a DVD player, and three other country-specific items (some items in ST26); and their responses on the number of cellular phones, televisions, computers, cars and the number of rooms with a bath or shower (ST27).

Home educational resources

The *index of home educational resources* (HEDRES) is based on the items measuring the existence of educational resources at home including a desk and a quiet place to study, a computer that students can use for schoolwork, educational software, books to help with students' school work, technical reference books and a dictionary (some items in ST26).

Cultural possessions

The *index of cultural possessions* (CULTPOSS) is based on students' responses to whether they had the following at home: classic literature, books of poetry and works of art (some items in ST26).

The rotated design of the student questionnaire

A major innovation in PISA 2012 is the rotated design of the student questionnaire. One of the main reasons for a rotated design, which had previously been implemented for the cognitive assessment, was to extend the content coverage of the student questionnaire. Table A1.3 provides an overview of the rotation design and content of questionnaire forms for the main survey.

The *PISA 2012 Technical Report* (OECD, forthcoming) provides all details regarding the rotated design of the student questionnaire in PISA 2012, including its implications in terms of *i*) proficiency estimates, *ii*) international reports and trends, *iii*) further analyses, *iv*) structure and documentation of the international database, and *v*) logistics. The rotated design has negligible implications for proficiency estimates and correlations of proficiency estimates with context constructs. The international database (available at www.pisa.oecd.org) includes all background variables for each student. The variables based on the questions that students answered reflect their responses; those that are based on questions that were not administered show a distinctive missing code. Rotation allows the estimation of a full co-variance matrix which means that all variables can be correlated with all other variables. It does not affect conclusions in terms of whether or not an effect would be considered significant in multilevel models.

School-level simple indices

School and class size

The *index of school size* (SCHSIZE) was derived by summing up the number of girls and boys at a school (SC07).

Student-teacher ratio

The *student-teacher ratio* (STRATIO) was obtained by dividing the school size by the total number of teachers (SC09). The number of part-time teachers was weighted by 0.5 and the number of full-time teachers was weighted by 1.0 in the computation of this index.



The *student-mathematics teacher ratio* (SMRATIO) was obtained by dividing the school size by the total number of mathematics teachers (SC10Q11 and SC10Q12). The number of part-time mathematics teachers was weighted by 0.5 and the number of full time mathematics teachers was weighted by 1.0 in the computation of this index.

School type

Schools are classified as either public or private, according to whether a private entity or a public agency has the ultimate power to make decisions concerning its affairs (SC01). This information is combined with SC02 which provides information on the percentage of total funding which comes from government sources to create the *index of school type* (SCHLTYPE). This index has three categories: (1) government-independent private schools controlled by a non-government organisation or with a governing board not selected by a government agency that receive less than 50% of their core funding from government agencies, (2) government-dependent private schools controlled by a non-government organisation or with a governing board not selected by a government agency that receive more than 50% of their core funding from government agencies, and (3) public schools controlled and managed by a public education authority or agency.

Availability of computers

The *index of computer availability* (RATCMP15) was derived from dividing the number of computers available for educational purposes available to students in the modal grade for 15-year-olds (SC11Q02) by the number of students in the modal grade for 15-year-olds (SC11Q01). The wording of the questions asking about computer availability changed between 2006 and 2009. Comparisons involving availability of computers are possible for 2012 data with 2009 data, but not with 2006 or earlier.

The *index of computers connected to the Internet* (COMPWEB) was derived from dividing the number of computers for educational purposes available to students in the modal grade for 15-year-olds that are connected to the web (SC11Q03) by the number of computers for educational purposes available to students in the modal grade for 15-year-olds (SC11Q02).

Quantity of teaching staff at school

The *proportion of fully certified teachers* (PROPCERT) was computed by dividing the number of fully certified teachers (SC09Q21 plus $0.5 \times \text{SC09Q22}$) by the total number of teachers (SC09Q11 plus $0.5 \times \text{SC09Q12}$). The proportion of teachers who have an ISCED 5A qualification (PROPQUAL) was calculated by dividing the number of these kind of teachers (SC09Q31 plus $0.5 \times \text{SC09Q32}$) by the total number of teachers (SC09Q11 plus $0.5 \times \text{SC09Q12}$). The proportion of mathematics teachers (PROPMATH) was computed by dividing the number of mathematics teachers (SC10Q11 plus $0.5 \times \text{SC10Q12}$) by the total number of teachers (SC09Q11 plus $0.5 \times \text{SC09Q12}$). The proportion of mathematics teachers who have an ISCED 5A qualification (PROPMA5A) was computed by dividing the number of mathematics teachers who have an ISCED 5A qualification (SC10Q21 plus $0.5 \times \text{SC10Q22}$) by the number of mathematics teachers (SC10Q11 plus $0.5 \times \text{SC10Q12}$).

Although both PISA 2003 and PISA 2012 asked school principals about the school's teaching staff, the wording of the questions on the proportion of teachers with an ISCED 5A qualification changed, rendering comparisons impossible.

Academic selectivity

The index of academic selectivity (SCHSEL) was derived from school principals' responses on how frequently consideration was given to the following two factors when students were admitted to the school, based on a scale with response categories "never", "sometimes" and "always" (SC32Q02 and SC32Q03): students' record of academic performance (including placement tests); and recommendation of feeder schools. This index has the following three categories: (1) schools where these two factors are "never" considered for admission, (2) schools considering at least one of these two factors "sometimes" but neither factor "always", and (3) schools where at least one of these two factors is "always" considered for admission.

Although both PISA 2003 and PISA 2012 asked school principals about the school's criteria for admitting students, the wording of the questions changed, rendering comparisons impossible.

Ability grouping

The *index of ability grouping in mathematics classes* (ABGMATH) was derived from the two items of school principals' reports on whether their school organises mathematics instruction differently for student with different abilities "for all classes", "for some classes", or "not for any classes" (SC15Q01 for mathematics classes study similar content but at different levels and SC15Q02 for different classes study different content or sets of mathematics topics that have different levels of difficulty). This index has the following three categories: (1) no mathematics classes study different levels of difficulty or different content (i.e. "not for any classes" for both SC15Q01 and SC15Q02); (2) some mathematics classes study different levels of difficulty or different content (i.e. "for some classes" for either SC15Q01 or SC15Q02); (3) all mathematics classes study different levels of difficulty or different content (i.e. "for all classes" for either SC15Q01 or SC15Q02).

Extracurricular activities offered by school

The *index of mathematics extracurricular activities at school* (MACTIV) was derived from school principals' reports on whether their schools offered the following activities to students in the national modal grade for 15-year-olds in the academic year of the PISA assessment (SC16 and SC21 for the last one): *i*) mathematics club, *ii*) mathematics competition, *iii*) club with a focus on computers/



Information, Communication Technology, and *iv*) additional mathematics lessons. This index was developed by summing up the number of activities that a school offers. For “additional mathematics lessons” (SC21), it is counted as one when school principals responded “enrichment mathematics only”, “remedial mathematics only” or “without differentiation depending on the prior achievement level of the students”; and it is counted as two when school principals responded “both enrichment and remedial mathematics”.

The *index of creative extracurricular activities at school* (CREACTIV) was derived from school principals’ reports on whether their schools offered the following activities to students in the national modal grade for 15-year-olds in the academic year of the PISA assessment (SC16): *i*) band, orchestra or choir, *ii*) school play or school musical, and *iii*) art club or art activities. This index was developed by adding up the number of activities that a school offers.

Use of assessment

School principals were asked to report whether students’ assessments are used for the following purposes (SC18): *i*) to inform parents about their child’s progress; *ii*) to make decisions about students’ retention or promotion; *iii*) to group students for instructional purposes; *iv*) to compare the school to district or national performance; *v*) to monitor the school’s progress from year to year; *vi*) to make judgements about teachers’ effectiveness; *vii*) to identify aspects of instruction or the curriculum that could be improved; and *viii*) to compare the school with other schools. The *index of use of assessment* (ASSESS) was derived from these eight items by adding up the number of “yes” in principals’ responses to these questions.

School responsibility for resource allocation

School principals were asked to report whether “principals”, “teachers”, “school governing board”, “regional or local education authority” or “national education authority” have a considerable responsibility for the following tasks (SC33): *i*) selecting teachers for hire; *ii*) firing teachers; *iii*) establishing teachers’ starting salaries; *iv*) determining teachers’ salary increases; *v*) formulating the school budget; and *vi*) deciding on budget allocations within the school. The *index of school responsibility for resource allocation* (RESPRES) was derived from these six items. The ratio of the number of responsibilities that “principals” and/or “teachers” have for these six items to the number of responsibilities that “regional or local education authority” and/or “national education authority” have for these six items was computed. Positive values on this index indicate relatively more responsibility for schools than local, regional or national education authority. This index has an OECD mean of 0 and a standard deviation of 1.

Although both PISA 2003 and PISA 2012 asked school principals about the school’s responsibility for resource allocation, the wording of the questions changed, rendering comparisons impossible.

School responsibility for curriculum and assessment

School principals were asked to report whether “principals”, “teachers”, “school governing board”, “regional or local education authority”, or “national education authority” have a considerable responsibility for the following tasks (SC33): *i*) establishing student assessment policies; *ii*) choosing which textbooks are used; *iii*) determining course content; and *iv*) deciding which courses are offered. The *index of the school responsibility for curriculum and assessment* (RESPCUR) was derived from these four items. The ratio of the number of responsibilities that “principals” and/or “teachers” have for these four items to the number of responsibilities that “regional or local education authority” and/or “national education authority” have for these four items was computed. Positive values on this index indicate relatively more responsibility for schools than local, regional or national education authority. This index has an OECD mean of 0 and a standard deviation of 1.

Although both PISA 2003 and PISA 2012 asked school principals about the school’s responsibility for admission and instruction policies, the wording of the questions changed, rendering comparisons impossible.

School-level scale indices

School principals’ leadership

The *index of school management: framing and communicating the school’s goals and curricular development* (LEADCOM) was derived from school principals’ responses about the frequency with which they were involved in the following school affairs in the previous school year (SC34): *i*) use student performance results to develop the school’s educational goals; *ii*) make sure that the professional development activities of teachers are in accordance with the teaching goals of the school; *iii*) ensure that teachers work according to the school’s educational goals; and *iv*) discuss the school’s academic goals with teachers at faculty meetings. The *index of school management: instructional leadership* (LEADINST) was derived from school principals’ responses about the frequency with which they were involved in the following school affairs in the previous school year (SC34): *i*) promote teaching practices based on recent educational research, *ii*) praise teachers whose students are actively participating in learning, and *iii*) draw teachers’ attention to the importance of pupils’ development of critical and social capacities. The *index of school management: promoting instructional improvements and professional development* (LEADPD) was derived from school principals’ responses about the frequency with which they were involved in the following school affairs in the previous school year (SC34): *i*) take the initiative to discuss matters, when a teacher has problems in his/her classroom; *ii*) pay attention to disruptive behaviour in classrooms; and *iii*) solve a problem together with a teacher, when the teacher brings up a classroom problem. The *index of school management: teacher participation* (LEADTCH) was derived from school principals’ responses about the frequency with which they were involved in the following school affairs in the previous school year (SC34): *i*) provide staff with opportunities to participate in school decision-making; *ii*) engage teachers to help build a school culture of continuous improvement; and *iii*) ask teachers to participate in reviewing management practices. Higher



values on these indices indicate greater involvement of school principals in school affairs.

Teacher shortage

The *index of teacher shortage* (TCSHORT) was derived from four items measuring school principals' perceptions of potential factors hindering instruction at their school (SC14). These factors are a lack of: *i*) qualified science teachers; *ii*) qualified mathematics teachers; *iii*) qualified <test language> teachers; and *iv*) qualified teachers of other subjects. Higher values on this index indicate school principals' reports of higher teacher shortage at a school.

For trends analyses, the PISA 2003 values of the index of teacher shortage were rescaled to be comparable to those in PISA 2012. As a result, values for the index of teacher shortage for PISA 2003 reported in this volume may differ from those reported in *Learning for Tomorrow's World: First Results from PISA 2003* (OECD, 2004).

Quality of school's educational resources

The *index of quality of school educational resources* (SCMATEDU) was derived from six items measuring school principals' perceptions of potential factors hindering instruction at their school (SC14). These factors are: *i*) shortage or inadequacy of science laboratory equipment; *ii*) shortage or inadequacy of instructional materials; *iii*) shortage or inadequacy of computers for instruction; *iv*) lack or inadequacy of Internet connectivity; *v*) shortage or inadequacy of computer software for instruction; and *vi*) shortage or inadequacy of library materials. As all items were inverted for scaling, higher values on this index indicate better quality of educational resources.

For trends analyses, the PISA 2003 values of the index of quality of educational resources were rescaled to be comparable to those in PISA 2012. As a result, values for the index of quality educational resources for PISA 2003 reported in this volume may differ from those reported in *Learning for Tomorrow's World: First Results from PISA 2003* (OECD, 2004). One of the questions included to compute the index of quality of educational resources in PISA 2012 ("lack or inadequacy of internet connection") was not included in the PISA 2003 questionnaire. Estimation of the PISA 2003 index treats this question as missing and, under the assumption that the relationship between the items remains unchanged with the inclusion of the new questions, the PISA 2003 and PISA 2012 values on the index of quality of educational resources are comparable after the rescaling.

Quality of schools' physical infrastructure

The *index of quality of physical infrastructure* (SCMATBUI) was derived from three items measuring school principals' perceptions of potential factors hindering instruction at their school (SC14). These factors are: *i*) shortage or inadequacy of school buildings and grounds; *ii*) shortage or inadequacy of heating/cooling and lighting systems; and *iii*) shortage or inadequacy of instructional space (e.g. classrooms). As all items were inverted for scaling, higher values on this index indicate better quality of physical infrastructure.

For trends analyses, the PISA 2003 values of the index of quality of physical infrastructure were rescaled to be comparable to those in PISA 2012. As a result, values for the index of quality of physical infrastructure for PISA 2003 reported in this volume may differ from those reported in *Learning for Tomorrow's World: First Results from PISA 2003* (OECD, 2004).

Teacher behaviour

The *index on teacher-related factors affecting school climate* (TEACCLIM) was derived from school principals' reports on the extent to which the learning of students was hindered by the following factors in their schools (SC22): *i*) students not being encouraged to achieve their full potential; *ii*) poor student-teacher relations; *iii*) teachers having to teach students of heterogeneous ability levels within the same class; *iv*) teachers having to teach students of diverse ethnic backgrounds (i.e. language, culture) within the same class; *v*) teachers' low expectations of students; *vi*) teachers not meeting individual students' needs; *vii*) teacher absenteeism; *viii*) staff resisting change; *ix*) teachers being too strict with students; *x*) teachers being late for classes; and *xi*) teachers not being well prepared for classes. As all items were inverted for scaling, higher values on this index indicate a positive teacher behaviour.

For trends analyses, the PISA 2003 values of the index of teacher-related factors affecting school climate were rescaled to be comparable to those in PISA 2012. As a result, values for the index of teacher-related factors affecting school climate for PISA 2003 reported in this volume may differ from those reported in *Learning for Tomorrow's World: First Results from PISA 2003* (OECD, 2004). Four of the questions included to compute the index of teacher-related factors affecting school climate in PISA 2012 ("teachers having to teach students of heterogeneous ability levels within the same class," "teachers having to teach students of diverse ethnic backgrounds (i.e. language, culture) within the same class," "teachers being late for classes," and "teachers not being well prepared for classes") were not included in the PISA 2003 questionnaire. Estimation of the PISA 2003 index treats these indices as missing and, under the assumption that the relationship between the items remains unchanged with the inclusion of the new questions, the PISA 2003 and PISA 2012 values on the index of teacher-related factors affecting school climate are comparable after the rescaling.

Student behaviour

The *index of student-related factors affecting school climate* (STUDCLIM) was derived from school principals' reports on the extent to which the learning of students was hindered by the following factors in their schools (SC22): *i*) student truancy; *ii*) students skipping classes; *iii*) students arriving late for school; *iv*) students not attending compulsory school events (e.g. sports day) or excursions, *v*) students lacking respect for teachers; *vi*) disruption of classes by students; *vii*) student use of alcohol or illegal drugs; and *viii*) students



intimidating or bullying other students. As all items were inverted for scaling, higher values on this index indicate a positive student behaviour.

For trends analyses, the PISA 2003 values of the index of student-related factors affecting school climate were rescaled to be comparable to those in PISA 2012. As a result, values for the index of student-related factors affecting school climate for PISA 2003 reported in this volume may differ from those reported in *Learning for Tomorrow's World: First Results from PISA 2003* (OECD, 2004). Two of the questions included to compute the index of student-related factors affecting school climate in PISA 2012 ("students arriving late for school," and "students not attending compulsory school events (e.g. sports day) or excursions") were not included in the PISA 2003 questionnaire. Estimation of the PISA 2003 index treats these questions as missing and, under the assumption that the relationship between the items remains unchanged with the inclusion of the new questions, the PISA 2003 and PISA 2012 values on the index of student-related factors affecting school climate are comparable after the rescaling.

Teacher morale

The *index of teacher morale* (TCMORALE) was derived from school principals' reports on the extent to which they agree with the following statements considering teachers in their schools (SC26): *i*) the morale of teachers in this school is high; *ii*) teachers work with enthusiasm; *iii*) teachers take pride in this school; and *iv*) teachers value academic achievement. As all items were inverted for scaling, higher values on this index indicate more positive teacher morale.

For trends analyses, the PISA 2003 values of the index of teacher morale were rescaled to be comparable to those in PISA 2012. As a result, values for the index teacher morale for PISA 2003 reported in this volume may differ from those reported in *Learning for Tomorrow's World: First Results from PISA 2003* (OECD, 2004).

Notes

- Note that for ISCO coding 0 "Arm forces", the following recoding was followed: "Officers" were coded as "Managers" (ISCO 1), and "Other armed forces occupations" (drivers, gunners, seaman, generic armed forces) as "Plant and Machine operators" (ISCO 8). In addition, all answers starting with "97" (housewives, students, and "vague occupations") were coded into missing.
- The update from ISCO-88 to ISCO-08 mainly involved *i*) more adequate categories for IT-related occupations, *ii*) distinction of military ranks and *iii*) a revision of the categories classifying different managers
- Information on ISCO08 and ISEI08 is included from <http://www.ilo.org/public/english/bureau/stat/isco/index.htm> and <http://home.fsw.vu.nl/hbg.ganzeboom/isco08>

References

- Ganzeboom, H.B.G. (2010), "A new international socio-economic index [ISEI] of occupational status for the International Standard Classification of Occupation 2008 [ISCO-08] constructed with data from the ISSP 2002-2007; with an analysis of quality of occupational measurement in ISSP ", paper presented at Annual Conference of International Social Survey Programme, Lisbon, 1 May 2010.
- Ganzeboom, H.B.G. and D.J. Treiman (2003), "Three Internationally Standardised Measures for Comparative Research on Occupational Status ", in Jürgen H.P. Hoffmeyer-Zlotnik and Christof Wolf (eds.), *Advances in Cross-National Comparison: A European Working Book for Demographic and Socio-Economic Variables*, Kluwer Academic Press, New York.
- Ganzeboom, H.B.G. and D.J. Treiman (1996), "Internationally Comparable Measures of Occupational Status for the 1988 International Standard Classification of Occupations", *Social Science Research*, Vol. 25, pp. 201-39.
- Ganzeboom, H.B.G., P. de Graaf and D.J. Treiman (1992), "A Standard International Socio-Economic Index of Occupational Status", *Social Science Research*, Vol. 21, Issue 1, pp. 1-56.
- Ganzeboom, H.B.G., R. Luijkx and D.J. Treiman (1989), "InterGenerational Class Mobility in Comparative Perspective", *Research in Social Stratification and Mobility*, Vol. 8, pp. 3-79.
- ILO (1990), *ISCO-88: International Standard Classification of Occupations*, International Labour Office, Geneva.
- OECD (forthcoming), *PISA 2012 Technical Report*, OECD Publishing.
- OECD (2013a), *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy*, PISA, OECD Publishing.
<http://dx.doi.org/10.1787/9789264190511-en>
- OECD (2013b), *Education at a Glance 2013: OECD Indicators*, OECD Publishing.



<http://dx.doi.org/10.1787/eag-2013-en>

OECD (2013c), *OECD Skills Outlook 2013: First Results from the Survey of Adult Skills*, OECD Publishing.
<http://dx.doi.org/10.1787/9789264204256-en>

OECD (2004), *Learning for Tomorrow's World: First Results from PISA 2003*, PISA, OECD Publishing.
<http://dx.doi.org/10.1787/9789264006416-en>

OECD (1999), *Classifying Educational Programmes: Manual for ISCED-97 Implementation in OECD Countries*, OECD Publishing.
www.oecd.org/education/skills-beyond-school/1962350.pdf

Warm, T.A. (1989), "Weighted likelihood estimation of ability in item response theory", *Psychometrika*, Volume 54, Issue 3, pp. 427-450.
<http://dx.doi.org/10.1007/BF02294627>



ANNEX A2

THE PISA TARGET POPULATION, THE PISA SAMPLES AND THE DEFINITION OF SCHOOLS

Definition of the PISA target population

PISA 2012 provides an assessment of the cumulative yield of education and learning at a point at which most young adults are still enrolled in initial education.

A major challenge for an international survey is to ensure that international comparability of national target populations is guaranteed in such a venture.

Differences between countries in the nature and extent of pre-primary education and care, the age of entry into formal schooling and the institutional structure of education systems do not allow the definition of internationally comparable grade levels of schooling. Consequently, international comparisons of education performance typically define their populations with reference to a target age group. Some previous international assessments have defined their target population on the basis of the grade level that provides maximum coverage of a particular age cohort. A disadvantage of this approach is that slight variations in the age distribution of students across grade levels often lead to the selection of different target grades in different countries, or between education systems within countries, raising serious questions about the comparability of results across, and at times within, countries. In addition, because not all students of the desired age are usually represented in grade-based samples, there may be a more serious potential bias in the results if the unrepresented students are typically enrolled in the next higher grade in some countries and the next lower grade in others. This would exclude students with potentially higher levels of performance in the former countries and students with potentially lower levels of performance in the latter.

In order to address this problem, PISA uses an age-based definition for its target population, i.e. a definition that is not tied to the institutional structures of national education systems. PISA assesses students who were aged between 15 years and 3 (complete) months and 16 years and 2 (complete) months at the beginning of the assessment period, plus or minus a 1 month allowable variation, and who were enrolled in an educational institution with Grade 7 or higher, regardless of the grade levels or type of institution in which they were enrolled, and regardless of whether they were in full-time or part-time education. Educational institutions are generally referred to as schools in this publication, although some educational institutions (in particular, some types of vocational education establishments) may not be termed schools in certain countries. As expected from this definition, the average age of students across OECD countries was 15 years and 9 months. The range in country means was 2 months and 5 days (0.18 years), from the minimum country mean of 15 years and 8 months to the maximum country mean of 15 years and 10 months.

Given this definition of population, PISA makes statements about the knowledge and skills of a group of individuals who were born within a comparable reference period, but who may have undergone different educational experiences both in and outside of schools. In PISA, these knowledge and skills are referred to as the yield of education at an age that is common across countries. Depending on countries' policies on school entry, selection and promotion, these students may be distributed over a narrower or a wider range of grades across different education systems, tracks or streams. It is important to consider these differences when comparing PISA results across countries, as observed differences between students at age 15 may no longer appear as students' educational experiences converge later on.

If a country's scale scores in reading, scientific or mathematical literacy are significantly higher than those in another country, it cannot automatically be inferred that the schools or particular parts of the education system in the first country are more effective than those in the second. However, one can legitimately conclude that the cumulative impact of learning experiences in the first country, starting in early childhood and up to the age of 15, and embracing experiences both in school, home and beyond, have resulted in higher outcomes in the literacy domains that PISA measures.

The PISA target population did not include residents attending schools in a foreign country. It does, however, include foreign nationals attending schools in the country of assessment.

To accommodate countries that desired grade-based results for the purpose of national analyses, PISA 2012 provided a sampling option to supplement age-based sampling with grade-based sampling.

Population coverage

All countries attempted to maximise the coverage of 15-year-olds enrolled in education in their national samples, including students enrolled in special educational institutions. As a result, PISA 2012 reached standards of population coverage that are unprecedented in international surveys of this kind.

The sampling standards used in PISA permitted countries to exclude up to a total of 5% of the relevant population either by excluding schools or by excluding students within schools. All but eight countries, Luxembourg (8.34%), Canada (6.37%), Denmark (6.10%), Norway (6.09%), Estonia (5.67%), Sweden (5.42%), the United Kingdom (5.36%) and the United States (5.34%), achieved this standard, and in 30 countries and economies, the overall exclusion rate was less than 2%. When language exclusions were accounted for (i.e. removed from the overall exclusion rate), Norway, Sweden, the United Kingdom and the United States no longer had an exclusion rate greater than 5%. For details, see www.pisa.oecd.org.



Exclusions within the above limits include:

- At the school level: *i*) schools that were geographically inaccessible or where the administration of the PISA assessment was not considered feasible; and *ii*) schools that provided teaching only for students in the categories defined under “within-school exclusions”, such as schools for the blind. The percentage of 15-year-olds enrolled in such schools had to be less than 2.5% of the nationally desired target population [0.5% maximum for *i*) and 2% maximum for *ii*)]. The magnitude, nature and justification of school-level exclusions are documented in the *PISA 2012 Technical Report* (OECD, forthcoming).
- At the student level: *i*) students with an intellectual disability; *ii*) students with a functional disability; *iii*) students with limited assessment language proficiency; *iv*) other – a category defined by the national centres and approved by the international centre; and *v*) students taught in a language of instruction for the main domain for which no materials were available. Students could not be excluded solely because of low proficiency or common discipline problems. The percentage of 15-year-olds excluded within schools had to be less than 2.5% of the nationally desired target population.

Table A2.1 describes the target population of the countries participating in PISA 2012. Further information on the target population and the implementation of PISA sampling standards can be found in the *PISA 2012 Technical Report* (OECD, forthcoming).

- **Column 1** shows the *total number of 15-year-olds* according to the most recent available information, which in most countries meant the year 2011 as the year before the assessment.
- **Column 2** shows the number of 15-year-olds enrolled in schools in Grade 7 or above (as defined above), which is referred to as the *eligible population*.
- **Column 3** shows the *national desired target population*. Countries were allowed to exclude up to 0.5% of students a priori from the eligible population, essentially for practical reasons. The following a priori exclusions exceed this limit but were agreed with the PISA Consortium: Belgium excluded 0.23% of its population for a particular type of student educated while working; Canada excluded 1.14% of its population from Territories and Aboriginal reserves; Chile excluded 0.04% of its students who live in Easter Island, Juan Fernandez Archipelago and Antarctica; Indonesia excluded 1.55% of its students from two provinces because of operational reasons; Ireland excluded 0.05% of its students in three island schools off the west coast; Latvia excluded 0.08% of its students in distance learning schools; and Serbia excluded 2.11% of its students taught in Serbian in Kosovo.
- **Column 4** shows the *number of students enrolled in schools that were excluded from the national desired target population* either from the sampling frame or later in the field during data collection.
- **Column 5** shows the *size of the national desired target population after subtracting the students enrolled in excluded schools*. This is obtained by subtracting Column 4 from Column 3.
- **Column 6** shows the *percentage of students enrolled in excluded schools*. This is obtained by dividing Column 4 by Column 3 and multiplying by 100.
- **Column 7** shows the *number of students participating in PISA 2012*. Note that in some cases this number does not account for 15-year-olds assessed as part of additional national options.
- **Column 8** shows the *weighted number of participating students*, i.e. the number of students in the nationally defined target population that the PISA sample represents.
- Each country attempted to maximise the coverage of the PISA target population within the sampled schools. In the case of each sampled school, all eligible students, namely those 15 years of age, regardless of grade, were first listed. Sampled students who were to be excluded had still to be included in the sampling documentation, and a list drawn up stating the reason for their exclusion. **Column 9** indicates the *total number of excluded students*, which is further described and classified into specific categories in Table A2.2.
- **Column 10** indicates the *weighted number of excluded students*, i.e. the overall number of students in the nationally defined target population represented by the number of students excluded from the sample, which is also described and classified by exclusion categories in Table A2.2. Excluded students were excluded based on five categories: *i*) students with an intellectual disability – the student has a mental or emotional disability and is cognitively delayed such that he/she cannot perform in the PISA testing situation; *ii*) students with a functional disability – the student has a moderate to severe permanent physical disability such that he/she cannot perform in the PISA testing situation; *iii*) students with a limited assessment language proficiency – the student is unable to read or speak any of the languages of the assessment in the country and would be unable to overcome the language barrier in the testing situation (typically a student who has received less than one year of instruction in the languages of the assessment may be excluded); *iv*) other – a category defined by the national centres and approved by the international centre; and *v*) students taught in a language of instruction for the main domain for which no materials were available.
- **Column 11** shows the *percentage of students excluded within schools*. This is calculated as the weighted number of excluded students (Column 10), divided by the weighted number of excluded and participating students (Column 8 plus Column 10), then multiplied by 100.
- **Column 12** shows the *overall exclusion rate*, which represents the weighted percentage of the national desired target population excluded from PISA either through school-level exclusions or through the exclusion of students within schools. It is calculated as the school-level exclusion rate (Column 6 divided by 100) plus within-school exclusion rate (Column 11 divided by 100) multiplied by 1 minus the school-level exclusion rate (Column 6 divided by 100). This result is then multiplied by 100.




[Part 1/2]
Table A2.1 PISA target populations and samples

		Population and sample information							
		Total population of 15-year-olds	Total enrolled population of 15-year-olds at Grade 7 or above	Total in national desired target population	Total school-level exclusions	Total in national desired target population after all school exclusions and before within-school exclusions	School-level exclusion rate (%)	Number of participating students	Weighted number of participating students
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OECD	Australia	291 967	288 159	288 159	5 702	282 457	1.98	17 774	250 779
	Austria	93 537	89 073	89 073	106	88 967	0.12	4 756	82 242
	Belgium	123 469	121 493	121 209	1 324	119 885	1.09	9 690	117 912
	Canada	417 873	409 453	404 767	2 936	401 831	0.73	21 548	348 070
	Chile	274 803	252 733	252 625	2 687	249 938	1.06	6 857	229 199
	Czech Republic	96 946	93 214	93 214	1 577	91 637	1.69	6 535	82 101
	Denmark	72 310	70 854	70 854	1 965	68 889	2.77	7 481	65 642
	Estonia	12 649	12 438	12 438	442	11 996	3.55	5 867	11 634
	Finland	62 523	62 195	62 195	523	61 672	0.84	8 829	60 047
	France	792 983	755 447	755 447	27 403	728 044	3.63	5 682	701 399
	Germany	798 136	798 136	798 136	10 914	787 222	1.37	5 001	756 907
	Greece	110 521	105 096	105 096	1 364	103 732	1.30	5 125	96 640
	Hungary	111 761	108 816	108 816	1 725	107 091	1.59	4 810	91 179
	Iceland	4 505	4 491	4 491	10	4 481	0.22	3 508	4 169
	Ireland	59 296	57 979	57 952	0	57 952	0.00	5 016	54 010
	Israel	118 953	113 278	113 278	2 784	110 494	2.46	6 061	107 745
	Italy	605 490	566 973	566 973	8 498	558 475	1.50	38 142	521 288
	Japan	1 241 786	1 214 756	1 214 756	26 099	1 188 657	2.15	6 351	1 128 179
	Korea	687 104	672 101	672 101	3 053	669 048	0.45	5 033	603 632
	Luxembourg	6 187	6 082	6 082	151	5 931	2.48	5 260	5 523
	Mexico	2 114 745	1 472 875	1 472 875	7 307	1 465 568	0.50	33 806	1 326 025
	Netherlands	194 000	193 190	193 190	7 546	185 644	3.91	4 460	196 262
	New Zealand	60 940	59 118	59 118	579	58 539	0.98	5 248	53 414
	Norway	64 917	64 777	64 777	750	64 027	1.16	4 686	59 432
	Poland	425 597	410 700	410 700	6 900	403 800	1.68	5 662	379 275
Portugal	108 728	127 537	127 537	0	127 537	0.00	5 722	96 034	
Slovak Republic	59 723	59 367	59 367	1 480	57 887	2.49	5 737	54 486	
Slovenia	19 471	18 935	18 935	115	18 820	0.61	7 229	18 303	
Spain	423 444	404 374	404 374	2 031	402 343	0.50	25 335	374 266	
Sweden	102 087	102 027	102 027	1 705	100 322	1.67	4 739	94 988	
Switzerland	87 200	85 239	85 239	2 479	82 760	2.91	11 234	79 679	
Turkey	1 266 638	965 736	965 736	10 387	955 349	1.08	4 848	866 681	
United Kingdom	738 066	745 581	745 581	19 820	725 761	2.66	12 659	688 236	
United States	3 985 714	4 074 457	4 074 457	41 142	4 033 315	1.01	6 111	3 536 153	
Partners	Albania	76 910	50 157	50 157	56	50 101	0.11	4 743	42 466
	Argentina	684 879	637 603	637 603	3 995	633 608	0.63	5 908	545 942
	Brazil	3 574 928	2 786 064	2 786 064	34 932	2 751 132	1.25	20 091	2 470 804
	Bulgaria	70 188	59 684	59 684	1 437	58 247	2.41	5 282	54 255
	Colombia	889 729	620 422	620 422	4	620 418	0.00	11 173	560 805
	Costa Rica	81 489	64 326	64 326	0	64 326	0.00	4 602	40 384
	Croatia	48 155	46 550	46 550	417	46 133	0.90	6 153	45 502
	Cyprus*	9 956	9 956	9 955	128	9 827	1.29	5 078	9 650
	Hong Kong-China	84 200	77 864	77 864	813	77 051	1.04	4 670	70 636
	Indonesia	4 174 217	3 599 844	3 544 028	8 039	3 535 989	0.23	5 622	2 645 155
	Jordan	129 492	125 333	125 333	141	125 192	0.11	7 038	111 098
	Kazakhstan	258 716	247 048	247 048	7 374	239 674	2.98	5 808	208 411
	Latvia	18 789	18 389	18 375	655	17 720	3.56	5 276	16 054
	Liechtenstein	417	383	383	1	382	0.26	293	314
	Lithuania	38 524	35 567	35 567	526	35 041	1.48	4 618	33 042
	Macao-China	6 600	5 416	5 416	6	5 410	0.11	5 335	5 366
	Malaysia	544 302	457 999	457 999	225	457 774	0.05	5 197	432 080
	Montenegro	8 600	8 600	8 600	18	8 582	0.21	4 744	7 714
	Peru	584 294	508 969	508 969	263	508 706	0.05	6 035	419 945
	Qatar	11 667	11 532	11 532	202	11 330	1.75	10 966	11 003
	Romania	146 243	146 243	146 243	5 091	141 152	3.48	5 074	140 915
	Russian Federation	1 272 632	1 268 814	1 268 814	17 800	1 251 014	1.40	6 418	1 172 539
	Serbia	80 089	75 870	74 272	1 987	72 285	2.67	4 684	67 934
	Shanghai-China	108 056	90 796	90 796	1 252	89 544	1.38	6 374	85 127
	Singapore	53 637	52 163	52 163	293	51 870	0.56	5 546	51 088
Chinese Taipei	328 356	328 336	328 336	1 747	326 589	0.53	6 046	292 542	
Thailand	982 080	784 897	784 897	9 123	775 774	1.16	6 606	703 012	
Tunisia	132 313	132 313	132 313	169	132 144	0.13	4 407	120 784	
United Arab Emirates	48 824	48 446	48 446	971	47 475	2.00	11 500	40 612	
Uruguay	54 638	46 442	46 442	14	46 428	0.03	5 315	39 771	
Viet Nam	1 717 996	1 091 462	1 091 462	7 729	1 083 733	0.71	4 959	956 517	

Notes: For a full explanation of the details in this table please refer to the PISA 2012 Technical Report (OECD, forthcoming). The figure for total national population of 15-year-olds enrolled in Column 2 may occasionally be larger than the total number of 15-year-olds in Column 1 due to differing data sources.

Information for the adjudicated regions is available on line.

* See notes at the beginning of this Annex.


StatLink  <http://dx.doi.org/10.1787/888932937092>

[Part 2/2]
Table A2.1 PISA target populations and samples

	Population and sample information				Coverage indices		
	Number of excluded students	Weighted number of excluded students	Within-school exclusion rate (%)	Overall exclusion rate (%)	Coverage index 1: Coverage of national desired population	Coverage index 2: Coverage of national enrolled population	Coverage index 3: Coverage of 15-year-old population
	(9)	(10)	(11)	(12)	(13)	(14)	(15)
OECD							
Australia	505	5 282	2.06	3.96	0.960	0.960	0.859
Austria	46	1 011	1.21	1.33	0.987	0.987	0.879
Belgium	39	367	0.31	1.39	0.986	0.984	0.955
Canada	1 796	21 013	5.69	6.37	0.936	0.926	0.833
Chile	18	548	0.24	1.29	0.987	0.987	0.834
Czech Republic	15	118	0.14	1.80	0.982	0.982	0.847
Denmark	368	2 381	3.50	6.10	0.938	0.938	0.908
Estonia	143	277	2.33	5.67	0.942	0.942	0.920
Finland	225	653	1.08	1.90	0.981	0.981	0.960
France	52	5 828	0.82	4.29	0.956	0.956	0.885
Germany	8	1 302	0.17	1.52	0.985	0.985	0.948
Greece	136	2 304	2.33	3.58	0.964	0.964	0.874
Hungary	27	928	1.01	2.55	0.974	0.974	0.816
Iceland	155	156	3.60	3.81	0.962	0.962	0.925
Ireland	271	2 524	4.47	4.47	0.955	0.955	0.911
Israel	114	1 884	1.72	4.07	0.959	0.959	0.906
Italy	741	9 855	1.86	3.30	0.967	0.967	0.861
Japan	0	0	0.00	2.10	0.979	0.979	0.909
Korea	17	2 238	0.37	0.82	0.992	0.992	0.879
Luxembourg	357	357	6.07	8.34	0.872	0.916	0.893
Mexico	58	3 247	0.24	0.74	0.993	0.993	0.627
Netherlands	27	1 056	0.54	4.27	0.956	0.956	1.012
New Zealand	255	2 030	3.66	4.60	0.954	0.954	0.876
Norway	278	3 133	5.01	6.09	0.939	0.939	0.916
Poland	212	11 566	2.96	4.56	0.954	0.954	0.891
Portugal	124	1 560	1.60	1.60	0.984	0.984	0.883
Slovak Republic	29	246	0.45	2.87	0.971	0.971	0.912
Slovenia	84	181	0.98	1.57	0.984	0.984	0.940
Spain	959	14 931	3.84	4.32	0.957	0.957	0.884
Sweden	201	3 789	3.84	5.42	0.946	0.946	0.930
Switzerland	256	1 093	1.35	4.14	0.958	0.958	0.914
Turkey	21	3 684	0.42	1.48	0.985	0.985	0.684
United Kingdom	486	20 173	2.85	5.36	0.946	0.946	0.932
United States	319	162 194	4.39	5.34	0.946	0.946	0.887
Partners							
Albania	1	10	0.02	0.13	0.999	0.999	0.552
Argentina	12	641	0.12	0.74	0.993	0.993	0.797
Brazil	44	4 900	0.20	1.43	0.986	0.986	0.691
Bulgaria	6	80	0.15	2.49	0.974	0.974	0.773
Colombia	23	789	0.14	0.14	0.999	0.999	0.630
Costa Rica	2	12	0.03	0.03	1.000	1.000	0.496
Croatia	91	627	1.36	2.23	0.978	0.978	0.945
Cyprus*	157	200	2.03	3.27	0.967	0.967	0.969
Hong Kong-China	38	518	0.73	1.75	0.982	0.982	0.839
Indonesia	2	860	0.03	0.26	0.997	0.982	0.634
Jordan	19	304	0.27	0.38	0.996	0.996	0.858
Kazakhstan	25	951	0.45	3.34	0.966	0.966	0.806
Latvia	14	76	0.47	3.89	0.960	0.959	0.854
Liechtenstein	13	13	3.97	4.22	0.958	0.958	0.753
Lithuania	130	867	2.56	3.98	0.960	0.960	0.858
Macao-China	3	3	0.06	0.17	0.998	0.998	0.813
Malaysia	7	554	0.13	0.18	0.998	0.998	0.794
Montenegro	4	8	0.10	0.31	0.997	0.997	0.897
Peru	8	549	0.13	0.18	0.998	0.998	0.719
Qatar	85	85	0.77	2.47	0.975	0.975	0.943
Romania	0	0	0.00	3.36	0.965	0.965	0.964
Russian Federation	69	11 940	1.01	2.38	0.976	0.976	0.921
Serbia	10	136	0.20	2.80	0.971	0.951	0.848
Shanghai-China	8	107	0.13	1.48	0.985	0.985	0.788
Singapore	33	315	0.61	1.17	0.988	0.988	0.952
Chinese Taipei	44	2 029	0.69	1.21	0.988	0.988	0.891
Thailand	12	1 144	0.16	1.31	0.987	0.987	0.716
Tunisia	5	130	0.11	0.24	0.998	0.998	0.913
United Arab Emirates	11	37	0.09	2.05	0.979	0.979	0.832
Uruguay	15	99	0.25	0.28	0.997	0.997	0.728
Viet Nam	1	198	0.02	0.72	0.993	0.993	0.557

Notes: For a full explanation of the details in this table please refer to the *PISA 2012 Technical Report* (OECD, forthcoming). The figure for total national population of 15-year-olds enrolled in Column 2 may occasionally be larger than the total number of 15-year-olds in Column 1 due to differing data sources. Information for the adjudicated regions is available on line.

* See notes at the beginning of this Annex.

StatLink  <http://dx.doi.org/10.1787/888932937092>



[Part 1/1]
Table A2.2 Exclusions

	Student exclusions (unweighted)						Student exclusions (weighted)					
	Number of excluded students with functional disability (Code 1)	Number of excluded students with intellectual disability (Code 2)	Number of excluded students because of language (Code 3)	Number of excluded students for other reasons (Code 4)	Number of excluded students because of no materials available in the language of instruction (Code 5)	Total number of excluded students	Weighted number of excluded students with functional disability (Code 1)	Weighted number of excluded students with intellectual disability (Code 2)	Weighted number of excluded students because of language (Code 3)	Weighted number of excluded students for other reasons (Code 4)	Number of excluded students because of no materials available in the language of instruction (Code 5)	Total weighted number of excluded students
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
OECD												
Australia	39	395	71	0	0	505	471	3 925	886	0	0	5 282
Austria	11	24	11	0	0	46	332	438	241	0	0	1 011
Belgium	5	22	12	0	0	39	24	154	189	0	0	367
Canada	82	1 593	121	0	0	1 796	981	18 682	1 350	0	0	21 013
Chile	3	15	0	0	0	18	74	474	0	0	0	548
Czech Republic	1	8	6	0	0	15	1	84	34	0	0	118
Denmark	10	204	112	42	0	368	44	1 469	559	310	0	2 381
Estonia	7	134	2	0	0	143	14	260	3	0	0	277
Finland	5	80	101	15	24	225	43	363	166	47	35	653
France	52	0	0	0	0	52	5 828	0	0	0	0	5 828
Germany	0	4	4	0	0	8	0	705	597	0	0	1 302
Greece	3	18	4	111	0	136	49	348	91	1 816	0	2 304
Hungary	1	15	2	9	0	27	36	568	27	296	0	928
Iceland	5	105	27	18	0	155	5	105	27	18	0	156
Ireland	13	159	33	66	0	271	121	1 521	283	599	0	2 524
Israel	9	91	14	0	0	114	133	1 492	260	0	0	1 884
Italy	64	566	111	0	0	741	596	7 899	1 361	0	0	9 855
Japan	0	0	0	0	0	0	0	0	0	0	0	0
Luxembourg	6	261	90	0	0	357	6	261	90	0	0	357
Mexico	21	36	1	0	0	58	812	2 390	45	0	0	3 247
Netherlands	5	21	1	0	0	27	188	819	50	0	0	1 056
New Zealand	27	118	99	0	11	255	235	926	813	0	57	2 030
Norway	11	192	75	0	0	278	120	2 180	832	0	0	3 133
Poland	23	89	6	88	6	212	1 470	5 187	177	4 644	89	11 566
Portugal	69	48	7	0	0	124	860	605	94	0	0	1 560
Korea	2	15	0	0	0	17	223	2 015	0	0	0	2 238
Slovak Republic	2	14	0	13	0	29	22	135	0	89	0	246
Slovenia	13	27	44	0	0	84	23	76	81	0	0	181
Spain	56	679	224	0	0	959	618	11 330	2 984	0	0	14 931
Sweden	120	0	81	0	0	201	2 218	0	1 571	0	0	3 789
Switzerland	7	99	150	0	0	256	41	346	706	0	0	1 093
Turkey	5	14	2	0	0	21	757	2 556	371	0	0	3 684
United Kingdom	40	405	41	0	0	486	1 468	15 514	3 191	0	0	20 173
United States	37	219	63	0	0	319	18 399	113 965	29 830	0	0	162 194
Partners												
Albania	0	0	1	0	0	1	0	0	10	0	0	10
Argentina	1	11	0	0	0	12	84	557	0	0	0	641
Brazil	17	27	0	0	0	44	1 792	3 108	0	0	0	4 900
Bulgaria	6	0	0	0	0	6	80	0	0	0	0	80
Colombia	12	10	1	0	0	23	397	378	14	0	0	789
Costa Rica	0	2	0	0	0	2	0	12	0	0	0	12
Croatia	10	78	3	0	0	91	69	539	19	0	0	627
Cyprus*	8	54	60	35	0	157	9	64	72	55	0	200
Hong Kong-China	4	33	1	0	0	38	57	446	15	0	0	518
Indonesia	1	0	1	0	0	2	426	0	434	0	0	860
Jordan	8	6	5	0	0	19	109	72	122	0	0	304
Kazakhstan	9	16	0	0	0	25	317	634	0	0	0	951
Latvia	3	7	4	0	0	14	8	45	24	0	0	76
Liechtenstein	1	7	5	0	0	13	1	7	5	0	0	13
Lithuania	10	120	0	0	0	130	66	801	0	0	0	867
Macao-China	0	1	2	0	0	3	0	1	2	0	0	3
Malaysia	3	4	0	0	0	7	274	279	0	0	0	554
Montenegro	3	1	0	0	0	4	7	1	0	0	0	8
Peru	3	5	0	0	0	8	269	280	0	0	0	549
Qatar	23	43	19	0	0	85	23	43	19	0	0	85
Romania	0	0	0	0	0	0	0	0	0	0	0	0
Russian Federation	25	40	4	0	0	69	4 345	6 934	660	0	0	11 940
Serbia	4	4	2	0	0	10	53	55	28	0	0	136
Shanghai-China	1	6	1	0	0	8	14	80	14	0	0	107
Singapore	5	17	11	0	0	33	50	157	109	0	0	315
Chinese Taipei	6	36	2	0	0	44	296	1 664	70	0	0	2 029
Thailand	2	10	0	0	0	12	13	1 131	0	0	0	1 144
Tunisia	4	1	0	0	0	5	104	26	0	0	0	130
United Arab Emirates	3	7	1	0	0	11	26	9	2	0	0	37
Uruguay	9	6	0	0	0	15	66	33	0	0	0	99
Viet Nam	0	1	0	0	0	1	0	198	0	0	0	198

Exclusion codes:

Code 1 Functional disability – student has a moderate to severe permanent physical disability.

Code 2 Intellectual disability – student has a mental or emotional disability and has either been tested as cognitively delayed or is considered in the professional opinion of qualified staff to be cognitively delayed.

Code 3 Limited assessment language proficiency – student is not a native speaker of any of the languages of the assessment in the country and has been resident in the country for less than one year.


Code 4 Other reasons defined by the national centres and approved by the international centre.

Code 5 No materials available in the language of instruction.

Note: For a full explanation of the details in this table please refer to the *PISA 2012 Technical Report* (OECD, forthcoming).

Information for the adjudicated regions is available on line.

* See notes at the beginning of this Annex.

StatLink  <http://dx.doi.org/10.1787/888932937092>



- **Column 13** presents an *index of the extent to which the national desired target population is covered by the PISA sample*. Canada, Denmark, Estonia, Luxembourg, Norway, Sweden, the United Kingdom and the United States were the only countries where the coverage is below 95%.
- **Column 14** presents an *index of the extent to which 15-year-olds enrolled in schools are covered by the PISA sample*. The index measures the overall proportion of the national enrolled population that is covered by the non-excluded portion of the student sample. The index takes into account both school-level and student-level exclusions. Values close to 100 indicate that the PISA sample represents the entire education system as defined for PISA 2012. The index is the weighted number of participating students (Column 8) divided by the weighted number of participating and excluded students (Column 8 plus Column 10), times the nationally defined target population (Column 5) divided by the eligible population (Column 2).
- **Column 15** presents an *index of the coverage of the 15-year-old population*. This index is the weighted number of participating students (Column 8) divided by the total population of 15-year-old students (Column 1).

This high level of coverage contributes to the comparability of the assessment results. For example, even assuming that the excluded students would have systematically scored worse than those who participated, and that this relationship is moderately strong, an exclusion rate in the order of 5% would likely lead to an overestimation of national mean scores of less than 5 score points (on a scale with an international mean of 500 score points and a standard deviation of 100 score points). This assessment is based on the following calculations: if the correlation between the propensity of exclusions and student performance is 0.3, resulting mean scores would likely be overestimated by 1 score point if the exclusion rate is 1%, by 3 score points if the exclusion rate is 5%, and by 6 score points if the exclusion rate is 10%. If the correlation between the propensity of exclusions and student performance is 0.5, resulting mean scores would be overestimated by 1 score point if the exclusion rate is 1%, by 5 score points if the exclusion rate is 5%, and by 10 score points if the exclusion rate is 10%. For this calculation, a model was employed that assumes a bivariate normal distribution for performance and the propensity to participate. For details, see the *PISA 2012 Technical Report* (OECD, forthcoming).

Sampling procedures and response rates

The accuracy of any survey results depends on the quality of the information on which national samples are based as well as on the sampling procedures. Quality standards, procedures, instruments and verification mechanisms were developed for PISA that ensured that national samples yielded comparable data and that the results could be compared with confidence.

Most PISA samples were designed as two-stage stratified samples (where countries applied different sampling designs, these are documented in the *PISA 2012 Technical Report* [OECD, forthcoming]). The first stage consisted of sampling individual schools in which 15-year-old students could be enrolled. Schools were sampled systematically with probabilities proportional to size, the measure of size being a function of the estimated number of eligible (15-year-old) students enrolled. A minimum of 150 schools were selected in each country (where this number existed), although the requirements for national analyses often required a somewhat larger sample. As the schools were sampled, replacement schools were simultaneously identified, in case a sampled school chose not to participate in PISA 2012.

In the case of Iceland, Liechtenstein, Luxembourg, Macao-China and Qatar, all schools and all eligible students within schools were included in the sample.

Experts from the PISA Consortium performed the sample selection process for most participating countries and monitored it closely in those countries that selected their own samples. The second stage of the selection process sampled students within sampled schools. Once schools were selected, a list of each sampled school's 15-year-old students was prepared. From this list, 35 students were then selected with equal probability (all 15-year-old students were selected if fewer than 35 were enrolled). The number of students to be sampled per school could deviate from 35, but could not be less than 20.

Data-quality standards in PISA required minimum participation rates for schools as well as for students. These standards were established to minimise the potential for response biases. In the case of countries meeting these standards, it was likely that any bias resulting from non-response would be negligible, i.e. typically smaller than the sampling error.

A minimum response rate of 85% was required for the schools initially selected. Where the initial response rate of schools was between 65% and 85%, however, an acceptable school response rate could still be achieved through the use of replacement schools. This procedure brought with it a risk of increased response bias. Participating countries were, therefore, encouraged to persuade as many of the schools in the original sample as possible to participate. Schools with a student participation rate between 25% and 50% were not regarded as participating schools, but data from these schools were included in the database and contributed to the various estimations. Data from schools with a student participation rate of less than 25% were excluded from the database.

PISA 2012 also required a minimum participation rate of 80% of students within participating schools. This minimum participation rate had to be met at the national level, not necessarily by each participating school. Follow-up sessions were required in schools in which too few students had participated in the original assessment sessions. Student participation rates were calculated over all original schools, and also over all schools, whether original sample or replacement schools, and from the participation of students in both the original assessment and any follow-up sessions. A student who participated in the original or follow-up cognitive sessions was regarded as a participant. Those who attended only the questionnaire session were included in the international database and contributed to the statistics presented in this publication if they provided at least a description of their father's or mother's occupation.




[Part 1/2]
Table A2.3 Response rates

	Initial sample – before school replacement					Final sample – after school replacement		
	Weighted school participation rate before replacement (%)	Weighted number of responding schools (weighted also by enrolment)	Weighted number of schools sampled (responding and non-responding) (weighted also by enrolment)	Number of responding schools (unweighted)	Number of non-responding and non-responding schools (unweighted)	Weighted school participation rate after replacement (%)	Weighted number of responding schools (weighted also by enrolment)	Weighted number of schools sampled (responding and non-responding) (weighted also by enrolment)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
OECD								
Australia	98	268 631	274 432	757	790	98	268 631	274 432
Austria	100	88 967	88 967	191	191	100	88 967	88 967
Belgium	84	100 482	119 019	246	294	97	115 004	119 006
Canada	91	362 178	396 757	828	907	93	368 600	396 757
Chile	92	220 009	239 429	200	224	99	236 576	239 370
Czech Republic	98	87 238	88 884	292	297	100	88 447	88 797
Denmark	87	61 749	71 015	311	366	96	67 709	70 892
Estonia	100	12 046	12 046	206	206	100	12 046	12 046
Finland	99	59 740	60 323	310	313	99	59 912	60 323
France	97	703 458	728 401	223	231	97	703 458	728 401
Germany	98	735 944	753 179	227	233	98	737 778	753 179
Greece	93	95 107	102 087	176	192	99	100 892	102 053
Hungary	98	99 317	101 751	198	208	99	101 187	101 751
Iceland	99	4 395	4 424	133	140	99	4 395	4 424
Ireland	99	56 962	57 711	182	185	99	57 316	57 711
Israel	91	99 543	109 326	166	186	94	103 075	109 895
Italy	89	478 317	536 921	1 104	1 232	97	522 686	536 821
Japan	86	1 015 198	1 175 794	173	200	96	1 123 211	1 175 794
Korea	100	661 575	662 510	156	157	100	661 575	662 510
Luxembourg	100	5 931	5 931	42	42	100	5 931	5 931
Mexico	92	1 323 816	1 442 242	1 431	1 562	95	1 374 615	1 442 234
Netherlands	75	139 709	185 468	148	199	89	165 635	185 320
New Zealand	81	47 441	58 676	156	197	89	52 360	58 616
Norway	85	54 201	63 653	177	208	95	60 270	63 642
Poland	85	343 344	402 116	159	188	98	393 872	402 116
Portugal	95	122 238	128 129	186	195	96	122 713	128 050
Slovak Republic	87	50 182	57 353	202	236	99	57 599	58 201
Slovenia	98	18 329	18 680	335	353	98	18 329	18 680
Spain	100	402 604	403 999	902	904	100	402 604	403 999
Sweden	99	98 645	99 726	207	211	100	99 536	99 767
Switzerland	94	78 825	83 450	397	422	98	82 032	83 424
Turkey	97	921 643	945 357	165	170	100	944 807	945 357
United Kingdom	80	564 438	705 011	477	550	89	624 499	699 839
United States	67	2 647 253	3 945 575	139	207	77	3 040 661	3 938 077
Partners								
Albania	100	49 632	49 632	204	204	100	49 632	49 632
Argentina	95	578 723	606 069	218	229	96	580 989	606 069
Brazil	93	2 545 863	2 745 045	803	886	95	2 622 293	2 747 688
Bulgaria	99	57 101	57 574	186	188	100	57 464	57 574
Colombia	87	530 553	612 605	323	363	97	596 557	612 261
Costa Rica	99	64 235	64 920	191	193	99	64 235	64 920
Croatia	99	45 037	45 636	161	164	100	45 608	45 636
Cyprus*	97	9 485	9 821	117	131	97	9 485	9 821
Hong Kong-China	79	60 277	76 589	123	156	94	72 064	76 567
Indonesia	95	2 799 943	2 950 696	199	210	98	2 892 365	2 951 028
Jordan	100	119 147	119 147	233	233	100	119 147	119 147
Kazakhstan	100	239 767	239 767	218	218	100	239 767	239 767
Latvia	88	15 371	17 488	186	213	100	17 428	17 448
Liechtenstein	100	382	382	12	12	100	382	382
Lithuania	98	33 989	34 614	211	216	100	34 604	34 604
Macao-China	100	5 410	5 410	45	45	100	5 410	5 410
Malaysia	100	455 543	455 543	164	164	100	455 543	455 543
Montenegro	100	8 540	8 540	51	51	100	8 540	8 540
Peru	98	503 915	514 574	238	243	99	507 602	514 574
Qatar	100	11 333	11 340	157	164	100	11 333	11 340
Romania	100	139 597	139 597	178	178	100	139 597	139 597
Russian Federation	100	1 243 564	1 243 564	227	227	100	1 243 564	1 243 564
Serbia	90	65 537	72 819	143	160	95	69 433	72 752
Shanghai-China	100	89 832	89 832	155	155	100	89 832	89 832
Singapore	98	50 415	51 687	170	176	98	50 945	51 896
Chinese Taipei	100	324 667	324 667	163	163	100	324 667	324 667
Thailand	98	757 516	772 654	235	240	100	772 452	772 654
Tunisia	99	129 229	130 141	152	153	99	129 229	130 141
United Arab Emirates	99	46 469	46 748	453	460	99	46 469	46 748
Uruguay	99	45 736	46 009	179	180	100	46 009	46 009
Viet Nam	100	1 068 462	1 068 462	162	162	100	1 068 462	1 068 462

Information for the adjudicated regions is available on line.

* See notes at the beginning of this Annex.

StatLink  <http://dx.doi.org/10.1787/888932937092>

[Part 2/2]

Table A2.3 Response rates

	Final sample – after school replacement		Final sample – students within schools after school replacement				
	Number of responding schools (unweighted)	Number of non-responding schools (unweighted)	Weighted student participation rate after replacement (%)	Number of students assessed (weighted)	Number of students sampled (assessed and absent) (weighted)	Number of students assessed (unweighted)	Number of students sampled (assessed and absent) (unweighted)
	(9)	(10)	(11)	(12)	(13)	(14)	(15)
OECD							
Australia	757	790	87	213 495	246 012	17 491	20 799
Austria	191	191	92	75 393	82 242	4 756	5 318
Belgium	282	294	91	103 914	114 360	9 649	10 595
Canada	840	907	81	261 928	324 328	20 994	25 835
Chile	221	224	95	214 558	226 689	6 857	7 246
Czech Republic	295	297	90	73 536	81 642	6 528	7 222
Denmark	339	366	89	56 096	62 988	7 463	8 496
Estonia	206	206	93	10 807	11 634	5 867	6 316
Finland	311	313	91	54 126	59 653	8 829	9 789
France	223	231	89	605 371	676 730	5 641	6 308
Germany	228	233	93	692 226	742 416	4 990	5 355
Greece	188	192	97	92 444	95 580	5 125	5 301
Hungary	204	208	93	84 032	90 652	4 810	5 184
Iceland	133	140	85	3 503	4 135	3 503	4 135
Ireland	183	185	84	45 115	53 644	5 016	5 977
Israel	172	186	90	91 181	101 288	6 061	6 727
Italy	1 186	1 232	93	473 104	510 005	38 084	41 003
Japan	191	200	96	1 034 803	1 076 786	6 351	6 609
Korea	156	157	99	595 461	603 004	5 033	5 101
Luxembourg	42	42	95	5 260	5 523	5 260	5 523
Mexico	1 468	1 562	94	1 193 866	1 271 639	33 786	35 972
Netherlands	177	199	85	148 432	174 697	4 434	5 215
New Zealand	177	197	85	40 397	47 703	5 248	6 206
Norway	197	208	91	51 155	56 286	4 686	5 156
Poland	182	188	88	325 389	371 434	5 629	6 452
Portugal	187	195	87	80 719	92 395	5 608	6 426
Slovak Republic	231	236	94	50 544	53 912	5 737	6 106
Slovenia	335	353	90	16 146	17 849	7 211	7 921
Spain	902	904	90	334 382	372 042	26 443	29 027
Sweden	209	211	92	87 359	94 784	4 739	5 141
Switzerland	410	422	92	72 116	78 424	11 218	12 138
Turkey	169	170	98	850 830	866 269	4 847	4 939
United Kingdom	505	550	86	528 231	613 736	12 638	14 649
United States	161	207	89	2 429 718	2 734 268	6 094	6 848
Partners							
Albania	204	204	92	39 275	42 466	4 743	5 102
Argentina	219	229	88	457 294	519 733	5 804	6 680
Brazil	837	886	90	2 133 035	2 368 438	19 877	22 326
Bulgaria	187	188	96	51 819	54 145	5 280	5 508
Colombia	352	363	93	507 178	544 862	11 164	12 045
Costa Rica	191	193	89	35 525	39 930	4 582	5 187
Croatia	163	164	92	41 912	45 473	6 153	6 675
Cyprus*	117	131	93	8 719	9 344	5 078	5 458
Hong Kong-China	147	156	93	62 059	66 665	4 659	5 004
Indonesia	206	210	95	2 478 961	2 605 254	5 579	5 885
Jordan	233	233	95	105 493	111 098	7 038	7 402
Kazakhstan	218	218	99	206 053	208 411	5 808	5 874
Latvia	211	213	91	14 579	16 039	5 276	5 785
Liechtenstein	12	12	93	293	314	293	314
Lithuania	216	216	92	30 429	33 042	4 618	5 018
Macao-China	45	45	99	5 335	5 366	5 335	5 366
Malaysia	164	164	94	405 983	432 080	5 197	5 529
Montenegro	51	51	94	7 233	7 714	4 799	5 117
Peru	240	243	96	398 193	414 728	6 035	6 291
Qatar	157	164	100	10 966	10 996	10 966	10 996
Romania	178	178	98	137 860	140 915	5 074	5 188
Russian Federation	227	227	97	1 141 317	1 172 539	6 418	6 602
Serbia	152	160	93	60 366	64 658	4 681	5 017
Shanghai-China	155	155	98	83 821	85 127	6 374	6 467
Singapore	172	176	94	47 465	50 330	5 546	5 887
Chinese Taipei	163	163	96	281 799	292 542	6 046	6 279
Thailand	239	240	99	695 088	702 818	6 606	6 681
Tunisia	152	153	90	108 342	119 917	4 391	4 857
United Arab Emirates	453	460	95	38 228	40 384	11 460	12 148
Uruguay	180	180	90	35 800	39 771	5 315	5 904
Viet Nam	162	162	100	955 222	956 517	4 959	4 966

Information for the adjudicated regions is available on line.

* See notes at the beginning of this Annex.


StatLink  <http://dx.doi.org/10.1787/888932937092>



Table A2.3 shows the response rates for students and schools, before and after replacement.

- **Column 1** shows the *weighted participation rate of schools before replacement*. This is obtained by dividing Column 2 by Column 3, multiply by 100.
- **Column 2** shows the *weighted number of responding schools before school replacement* (weighted by student enrolment).
- **Column 3** shows the *weighted number of sampled schools before school replacement* (including both responding and non-responding schools, weighted by student enrolment).
- **Column 4** shows the *unweighted number of responding schools before school replacement*.
- **Column 5** shows the *unweighted number of responding and non-responding schools before school replacement*.
- **Column 6** shows the *weighted participation rate of schools after replacement*. This is obtained by dividing Column 7 by Column 8, multiply by 100.
- **Column 7** shows the *weighted number of responding schools after school replacement* (weighted by student enrolment).
- **Column 8** shows the *weighted number of schools sampled after school replacement* (including both responding and non-responding schools, weighted by student enrolment).
- **Column 9** shows the *unweighted number of responding schools after school replacement*.
- **Column 10** shows the *unweighted number of responding and non-responding schools after school replacement*.
- **Column 11** shows the *weighted student participation rate after replacement*. This is obtained by dividing Column 12 by Column 13, multiply by 100.
- **Column 12** shows the *weighted number of students assessed*.
- **Column 13** shows the *weighted number of students sampled* (including both students who were assessed and students who were absent on the day of the assessment).
- **Column 14** shows the *unweighted number of students assessed*. Note that any students in schools with student-response rates less than 50% were not included in these rates (both weighted and unweighted).
- **Column 15** shows the *unweighted number of students sampled* (including both students that were assessed and students who were absent on the day of the assessment). Note that any students in schools where fewer than half of the eligible students were assessed were not included in these rates (neither weighted nor unweighted).

Definition of schools

In some countries, sub-units within schools were sampled instead of schools and this may affect the estimation of the between-school variance components. In Austria, the Czech Republic, Germany, Hungary, Japan, Romania and Slovenia, schools with more than one study programme were split into the units delivering these programmes. In the Netherlands, for schools with both lower and upper secondary programmes, schools were split into units delivering each programme level. In the Flemish community of Belgium, in the case of multi-campus schools, implantations (campuses) were sampled, whereas in the French Community, in the case of multi-campus schools, the larger administrative units were sampled. In Australia, for schools with more than one campus, the individual campuses were listed for sampling. In Argentina, Croatia and Dubai (United Arab Emirates), schools that had more than one campus had the locations listed for sampling. In Spain, the schools in the Basque region with multi-linguistic models were split into linguistic models for sampling.

Grade levels

Students assessed in PISA 2012 are at various grade levels. The percentage of students at each grade level is presented by country and economy in Table A2.4a and by gender within each country and economy in Table A2.4b.


[Part 1/1]

Table A2.4a Percentage of students at each grade level

	All students											
	7th grade		8th grade		9th grade		10th grade		11th grade		12th grade and above	
	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.
OECD												
Australia	0.0	(0.0)	0.1	(0.0)	10.8	(0.5)	70.0	(0.6)	19.1	(0.4)	0.0	(0.0)
Austria	0.3	(0.1)	5.4	(0.7)	43.3	(0.9)	51.0	(1.0)	0.1	(0.0)	0.0	c
Belgium	0.9	(0.1)	6.4	(0.5)	30.9	(0.6)	60.8	(0.6)	1.0	(0.1)	0.0	(0.0)
Canada	0.1	(0.0)	1.1	(0.1)	13.2	(0.6)	84.6	(0.6)	1.0	(0.1)	0.1	(0.0)
Chile	1.4	(0.3)	4.1	(0.6)	21.7	(0.8)	66.1	(1.2)	6.7	(0.3)	0.0	c
Czech Republic	0.4	(0.1)	4.5	(0.4)	51.1	(1.2)	44.1	(1.3)	0.0	c	0.0	c
Denmark	0.1	(0.0)	18.2	(0.8)	80.6	(0.8)	1.0	(0.2)	0.0	c	0.0	c
Estonia	0.6	(0.2)	22.1	(0.7)	75.4	(0.7)	1.9	(0.3)	0.0	c	0.0	c
Finland	0.7	(0.2)	14.2	(0.4)	85.0	(0.4)	0.0	c	0.1	(0.1)	0.0	c
France	0.0	(0.0)	1.9	(0.3)	27.9	(0.7)	66.6	(0.7)	3.5	(0.3)	0.1	(0.1)
Germany	0.6	(0.1)	10.0	(0.6)	51.9	(0.8)	36.7	(0.9)	0.8	(0.4)	0.0	c
Greece	0.3	(0.1)	1.2	(0.3)	4.0	(0.7)	94.5	(1.0)	0.0	c	0.0	c
Hungary	2.8	(0.5)	8.7	(0.9)	67.8	(0.9)	20.6	(0.6)	0.0	c	0.0	c
Iceland	0.0	c	0.0	c	0.0	c	100.0	c	0.0	c	0.0	c
Ireland	0.0	(0.0)	1.9	(0.2)	60.5	(0.8)	24.3	(1.2)	13.3	(1.0)	0.0	c
Israel	0.0	(0.0)	0.3	(0.1)	17.1	(0.9)	81.7	(0.9)	0.8	(0.3)	0.0	c
Italy	0.4	(0.1)	1.7	(0.2)	16.8	(0.6)	78.5	(0.7)	2.6	(0.2)	0.0	(0.0)
Japan	0.0	c	0.0	c	0.0	c	100.0	c	0.0	c	0.0	c
Korea	0.0	c	0.0	c	5.9	(0.8)	93.8	(0.8)	0.2	(0.1)	0.0	c
Luxembourg	0.7	(0.1)	10.2	(0.2)	50.7	(0.1)	38.0	(0.1)	0.5	(0.1)	0.0	c
Mexico	1.1	(0.1)	5.2	(0.3)	30.8	(1.0)	60.8	(1.1)	2.1	(0.3)	0.1	(0.0)
Netherlands	0.0	c	3.6	(0.4)	46.7	(1.0)	49.2	(1.1)	0.5	(0.1)	0.0	c
New Zealand	0.0	c	0.0	c	0.1	(0.1)	6.2	(0.4)	88.3	(0.5)	5.4	(0.4)
Norway	0.0	c	0.0	c	0.4	(0.1)	99.4	(0.1)	0.2	(0.0)	0.0	c
Poland	0.5	(0.1)	4.1	(0.4)	94.9	(0.4)	0.5	(0.2)	0.0	c	0.0	c
Portugal	2.4	(0.3)	8.2	(0.7)	28.6	(1.6)	60.5	(2.1)	0.3	(0.1)	0.0	c
Slovak Republic	1.7	(0.3)	4.5	(0.5)	39.5	(1.5)	52.7	(1.4)	1.6	(0.5)	0.0	c
Slovenia	0.0	c	0.3	(0.2)	5.1	(0.8)	90.7	(0.8)	3.9	(0.2)	0.0	c
Spain	0.1	(0.0)	9.8	(0.5)	24.1	(0.4)	66.0	(0.6)	0.0	(0.0)	0.0	c
Sweden	0.0	(0.0)	3.7	(0.3)	94.0	(0.6)	2.2	(0.5)	0.0	c	0.0	c
Switzerland	0.6	(0.1)	12.9	(0.8)	60.6	(1.0)	25.6	(1.0)	0.2	(0.1)	0.0	c
Turkey	0.5	(0.2)	2.2	(0.3)	27.6	(1.2)	65.5	(1.2)	4.0	(0.3)	0.3	(0.1)
United Kingdom	0.0	c	0.0	c	0.0	(0.0)	1.3	(0.3)	95.0	(0.3)	3.6	(0.1)
United States	0.0	c	0.3	(0.1)	11.7	(1.1)	71.2	(1.1)	16.6	(0.8)	0.2	(0.1)
OECD average	0.5	(0.0)	4.9	(0.1)	34.7	(0.1)	51.9	(0.2)	7.7	(0.1)	0.3	(0.0)
Partners												
Albania	0.1	(0.1)	2.2	(0.3)	39.4	(2.4)	58.0	(2.5)	0.3	(0.1)	0.0	c
Argentina	2.0	(0.5)	12.0	(1.2)	22.6	(1.4)	59.4	(2.1)	2.8	(0.6)	1.1	(0.7)
Brazil	0.0	c	6.9	(0.5)	13.5	(0.7)	34.9	(1.0)	42.0	(1.0)	2.6	(0.2)
Bulgaria	0.9	(0.2)	4.6	(0.5)	89.5	(0.7)	4.9	(0.4)	0.0	(0.0)	0.0	c
Colombia	5.5	(0.6)	12.1	(0.7)	21.5	(0.8)	40.2	(0.9)	20.7	(1.0)	0.0	c
Costa Rica	7.4	(0.9)	13.7	(0.9)	39.6	(1.3)	39.1	(1.8)	0.2	(0.1)	0.0	c
Croatia	0.0	c	0.0	c	79.8	(0.4)	20.2	(0.4)	0.0	c	0.0	c
Cyprus*	0.0	(0.0)	0.5	(0.1)	4.5	(0.1)	94.3	(0.1)	0.7	(0.0)	0.0	(0.0)
Hong Kong-China	1.1	(0.1)	6.5	(0.4)	25.9	(0.7)	65.0	(0.9)	1.5	(1.4)	0.0	c
Indonesia	1.9	(0.4)	8.3	(0.8)	37.7	(2.6)	47.7	(3.0)	3.9	(0.6)	0.6	(0.6)
Jordan	0.1	(0.0)	1.1	(0.1)	6.0	(0.4)	92.9	(0.4)	0.0	c	0.0	c
Kazakhstan	0.2	(0.1)	4.9	(0.5)	67.2	(1.9)	27.4	(2.0)	0.2	(0.1)	0.1	(0.1)
Latvia	2.1	(0.4)	14.8	(0.7)	80.0	(0.8)	3.0	(0.4)	0.0	(0.0)	0.0	c
Liechtenstein	4.9	(0.7)	14.2	(1.5)	66.3	(1.3)	14.6	(0.2)	0.0	c	0.0	c
Lithuania	0.2	(0.1)	6.2	(0.6)	81.2	(0.7)	12.4	(0.7)	0.0	(0.0)	0.0	c
Macao-China	5.4	(0.1)	16.4	(0.2)	33.2	(0.2)	44.6	(0.1)	0.4	(0.1)	0.0	(0.0)
Malaysia	0.0	c	0.1	(0.0)	4.0	(0.5)	96.0	(0.5)	0.0	(0.0)	0.0	c
Montenegro	0.0	c	0.1	(0.0)	79.5	(0.1)	20.4	(0.1)	0.0	c	0.0	c
Peru	2.7	(0.4)	7.8	(0.5)	18.1	(0.7)	47.7	(0.9)	23.7	(0.8)	0.0	c
Qatar	0.9	(0.0)	3.1	(0.1)	13.8	(0.1)	64.8	(0.1)	17.1	(0.1)	0.3	(0.0)
Romania	0.2	(0.1)	7.4	(0.5)	87.2	(0.6)	5.1	(0.4)	0.0	c	0.0	c
Russian Federation	0.6	(0.1)	8.1	(0.5)	73.8	(1.6)	17.4	(1.8)	0.1	(0.1)	0.0	c
Serbia	0.1	(0.1)	1.5	(0.7)	96.7	(0.7)	1.7	(0.2)	0.0	c	0.0	c
Shanghai-China	1.1	(0.2)	4.5	(0.6)	39.6	(1.5)	54.2	(1.3)	0.6	(0.1)	0.1	(0.1)
Singapore	0.4	(0.1)	2.0	(0.2)	8.0	(0.3)	89.6	(0.3)	0.1	(0.1)	0.0	c
Chinese Taipei	0.0	c	0.2	(0.1)	36.2	(0.7)	63.6	(0.7)	0.0	c	0.0	c
Thailand	0.1	(0.0)	0.3	(0.1)	20.7	(1.0)	76.0	(1.1)	2.9	(0.5)	0.0	c
Tunisia	5.0	(0.6)	11.8	(1.3)	20.6	(1.4)	56.7	(2.7)	5.9	(0.5)	0.0	c
United Arab Emirates	0.9	(0.2)	2.8	(0.2)	11.3	(0.8)	61.9	(1.0)	22.2	(0.7)	0.9	(0.2)
Uruguay	6.9	(0.8)	12.2	(0.6)	22.4	(1.0)	57.3	(1.5)	1.3	(0.2)	0.0	c
Viet Nam	0.4	(0.2)	2.7	(0.7)	8.3	(1.7)	88.6	(2.3)	0.0	c	0.0	c

Information for the adjudicated regions is available on line.

* See notes at the beginning of this Annex.

StatLink  <http://dx.doi.org/10.1787/888932937092>




[Part 1/2]
Table A2.4b Percentage of students at each grade level, by gender

	Boys											
	7th grade		8th grade		9th grade		10th grade		11th grade		12th grade and above	
	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.
OECD												
Australia	0.0	c	0.1	(0.0)	13.1	(0.9)	69.2	(0.9)	17.5	(0.6)	0.0	(0.0)
Austria	0.3	(0.1)	6.0	(0.9)	44.8	(1.4)	48.9	(1.5)	0.0	c	0.0	c
Belgium	1.0	(0.1)	7.1	(0.6)	33.8	(0.9)	57.1	(1.0)	1.0	(0.2)	0.0	(0.0)
Canada	0.1	(0.1)	1.3	(0.2)	14.8	(0.8)	82.7	(0.8)	0.9	(0.1)	0.1	(0.1)
Chile	1.4	(0.4)	5.0	(0.9)	24.2	(1.0)	63.1	(1.6)	6.4	(0.4)	0.0	c
Czech Republic	0.7	(0.2)	5.5	(0.6)	54.9	(2.0)	39.0	(2.1)	0.0	c	0.0	c
Denmark	0.1	(0.0)	23.4	(1.0)	75.7	(1.0)	0.8	(0.3)	0.0	c	0.0	c
Estonia	0.8	(0.3)	25.7	(1.0)	71.7	(1.1)	1.7	(0.4)	0.0	c	0.0	c
Finland	0.9	(0.4)	16.2	(0.6)	82.8	(0.7)	0.0	c	0.1	(0.1)	0.0	c
France	0.1	(0.1)	2.3	(0.4)	30.8	(0.9)	63.5	(1.0)	3.2	(0.5)	0.1	(0.1)
Germany	0.9	(0.2)	11.6	(0.7)	53.6	(1.1)	33.2	(1.2)	0.7	(0.3)	0.0	c
Greece	0.4	(0.2)	1.8	(0.6)	4.8	(1.0)	93.0	(1.4)	0.0	c	0.0	c
Hungary	3.9	(0.6)	12.1	(1.5)	67.1	(1.3)	17.0	(0.8)	0.0	c	0.0	c
Iceland	0.0	c	0.0	c	0.0	c	100.0	c	0.0	c	0.0	c
Ireland	0.0	c	2.4	(0.3)	63.6	(1.0)	21.1	(1.4)	13.0	(1.3)	0.0	c
Israel	0.1	(0.1)	0.3	(0.1)	18.9	(1.3)	79.6	(1.3)	1.2	(0.5)	0.0	c
Italy	0.5	(0.2)	2.1	(0.3)	19.3	(0.7)	75.8	(0.7)	2.3	(0.2)	0.0	c
Japan	0.0	c	0.0	c	0.0	c	100.0	c	0.0	c	0.0	c
Korea	0.0	c	0.0	c	6.4	(1.2)	93.4	(1.2)	0.2	(0.1)	0.0	c
Luxembourg	0.7	(0.1)	10.7	(0.2)	51.1	(0.2)	37.0	(0.2)	0.6	(0.1)	0.0	c
Mexico	1.3	(0.2)	6.3	(0.3)	33.0	(1.1)	57.2	(1.2)	2.1	(0.5)	0.0	(0.0)
Netherlands	0.0	c	4.4	(0.6)	49.5	(1.1)	45.7	(1.2)	0.4	(0.1)	0.0	c
New Zealand	0.0	c	0.0	c	0.2	(0.1)	7.0	(0.5)	88.0	(0.7)	4.8	(0.5)
Norway	0.0	c	0.0	c	0.6	(0.1)	99.1	(0.1)	0.3	(0.0)	0.0	c
Poland	0.9	(0.2)	5.7	(0.6)	93.0	(0.6)	0.4	(0.2)	0.0	c	0.0	c
Portugal	2.6	(0.5)	9.9	(0.9)	30.1	(1.7)	57.0	(2.2)	0.4	(0.2)	0.0	c
Slovak Republic	1.5	(0.3)	5.4	(0.8)	40.1	(2.0)	51.5	(2.1)	1.5	(0.5)	0.0	c
Slovenia	0.0	c	0.4	(0.3)	6.3	(1.0)	90.2	(1.0)	3.1	(0.4)	0.0	c
Spain	0.1	(0.1)	11.8	(0.6)	25.8	(0.6)	62.2	(0.7)	0.1	(0.1)	0.0	c
Sweden	0.1	(0.1)	4.6	(0.5)	93.7	(0.8)	1.7	(0.6)	0.0	c	0.0	c
Switzerland	0.5	(0.1)	13.9	(0.9)	60.6	(1.7)	24.7	(2.0)	0.2	(0.1)	0.0	c
Turkey	0.3	(0.1)	2.6	(0.5)	33.2	(1.5)	60.3	(1.5)	3.2	(0.4)	0.3	(0.1)
United Kingdom	0.0	c	0.0	c	0.0	(0.0)	1.7	(0.4)	94.7	(0.4)	3.7	(0.2)
United States	0.0	c	0.4	(0.2)	14.6	(1.1)	69.8	(1.1)	14.9	(0.9)	0.3	(0.2)
OECD average	0.6	(0.1)	5.9	(0.1)	35.6	(0.2)	50.1	(0.2)	7.5	(0.1)	0.3	(0.1)
Partners												
Albania	0.1	(0.1)	2.9	(0.4)	42.9	(2.7)	53.8	(2.8)	0.2	(0.1)	0.0	c
Argentina	2.8	(0.8)	15.0	(1.7)	25.8	(1.9)	52.6	(2.6)	3.0	(0.9)	0.8	(0.5)
Brazil	0.0	c	9.0	(0.7)	15.8	(0.8)	36.1	(1.1)	37.2	(1.0)	1.9	(0.2)
Bulgaria	1.3	(0.3)	5.8	(0.7)	88.2	(1.0)	4.6	(0.4)	0.0	c	0.0	c
Colombia	7.4	(0.8)	13.5	(1.0)	22.1	(1.0)	38.8	(1.4)	18.2	(1.2)	0.0	c
Costa Rica	9.3	(1.3)	16.4	(1.2)	38.5	(1.5)	35.7	(2.0)	0.0	(0.0)	0.0	c
Croatia	0.0	c	0.0	c	82.0	(0.6)	18.0	(0.6)	0.0	c	0.0	c
Cyprus*	0.0	(0.0)	0.5	(0.1)	4.7	(0.1)	94.0	(0.2)	0.7	(0.1)	0.0	c
Hong Kong-China	1.2	(0.2)	6.9	(0.5)	27.5	(0.7)	63.0	(1.0)	1.4	(1.3)	0.0	c
Indonesia	2.3	(0.4)	10.0	(1.1)	38.5	(3.0)	45.5	(3.7)	3.1	(0.6)	0.6	(0.6)
Jordan	0.1	(0.1)	0.8	(0.2)	5.7	(0.6)	93.4	(0.6)	0.0	c	0.0	c
Kazakhstan	0.3	(0.1)	5.5	(0.6)	68.4	(2.4)	25.4	(2.6)	0.2	(0.1)	0.2	(0.2)
Latvia	3.6	(0.8)	18.0	(0.9)	76.4	(1.3)	2.0	(0.3)	0.0	(0.0)	0.0	c
Liechtenstein	4.5	(1.2)	16.5	(2.1)	69.4	(2.2)	9.6	(0.6)	0.0	c	0.0	c
Lithuania	0.2	(0.1)	7.3	(0.6)	82.2	(0.9)	10.4	(0.8)	0.0	(0.0)	0.0	c
Macao-China	7.1	(0.2)	19.3	(0.2)	33.3	(0.2)	40.0	(0.2)	0.2	(0.1)	0.0	(0.0)
Malaysia	0.0	c	0.1	(0.1)	5.1	(0.7)	94.7	(0.7)	0.0	c	0.0	c
Montenegro	0.0	c	0.1	(0.1)	82.0	(0.3)	17.9	(0.3)	0.0	c	0.0	c
Peru	3.1	(0.5)	9.1	(0.8)	19.5	(0.7)	46.2	(1.0)	22.1	(0.9)	0.0	c
Qatar	1.2	(0.1)	3.6	(0.1)	14.0	(0.1)	64.6	(0.2)	16.1	(0.2)	0.4	(0.0)
Romania	0.3	(0.2)	6.5	(0.6)	88.7	(0.7)	4.5	(0.4)	0.0	c	0.0	c
Russian Federation	0.7	(0.2)	8.9	(0.7)	73.7	(1.5)	16.7	(1.8)	0.1	(0.1)	0.0	c
Serbia	0.1	(0.1)	1.9	(0.9)	96.7	(1.0)	1.4	(0.2)	0.0	c	0.0	c
Shanghai-China	1.3	(0.3)	5.3	(0.8)	41.6	(1.6)	51.2	(1.4)	0.6	(0.1)	0.0	(0.0)
Singapore	0.4	(0.1)	2.0	(0.3)	8.3	(0.4)	89.3	(0.5)	0.0	(0.0)	0.0	c
Chinese Taipei	0.0	c	0.2	(0.2)	37.4	(1.5)	62.4	(1.5)	0.0	c	0.0	c
Thailand	0.1	(0.1)	0.4	(0.2)	22.9	(1.3)	74.1	(1.5)	2.5	(0.5)	0.0	c
Tunisia	6.3	(0.8)	14.6	(1.6)	21.9	(1.6)	52.3	(3.0)	4.9	(0.5)	0.0	c
United Arab Emirates	1.3	(0.3)	3.1	(0.3)	12.9	(0.9)	60.3	(1.2)	21.8	(1.0)	0.6	(0.1)
Uruguay	9.4	(1.3)	13.1	(0.8)	24.0	(1.1)	52.4	(1.9)	1.2	(0.2)	0.0	c
Viet Nam	0.7	(0.3)	3.5	(0.8)	10.5	(2.2)	85.3	(2.8)	0.0	c	0.0	c

Information for the adjudicated regions is available on line.

* See notes at the beginning of this Annex.

StatLink  <http://dx.doi.org/10.1787/888932937092>


[Part 2/2]

Table A2.4b Percentage of students at each grade level, by gender

	Girls											
	7th grade		8th grade		9th grade		10th grade		11th grade		12th grade and above	
	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.
OECD												
Australia	0.0	(0.0)	0.2	(0.1)	8.3	(0.3)	70.8	(0.6)	20.7	(0.6)	0.0	(0.0)
Austria	0.3	(0.1)	4.7	(0.7)	41.8	(1.3)	53.1	(1.4)	0.1	(0.1)	0.0	c
Belgium	0.9	(0.1)	5.7	(0.5)	28.0	(0.7)	64.4	(0.8)	1.0	(0.2)	0.0	c
Canada	0.1	(0.0)	0.9	(0.1)	11.5	(0.5)	86.4	(0.5)	1.2	(0.1)	0.0	(0.0)
Chile	1.3	(0.3)	3.3	(0.6)	19.3	(1.0)	69.0	(1.2)	7.1	(0.4)	0.0	c
Czech Republic	0.1	(0.1)	3.5	(0.5)	47.1	(2.0)	49.4	(2.1)	0.0	c	0.0	c
Denmark	0.1	(0.0)	13.0	(0.9)	85.6	(0.9)	1.3	(0.3)	0.0	c	0.0	c
Estonia	0.3	(0.1)	18.6	(0.8)	79.0	(0.9)	2.2	(0.4)	0.0	c	0.0	c
Finland	0.5	(0.1)	12.0	(0.4)	87.3	(0.4)	0.0	c	0.2	(0.1)	0.0	c
France	0.0	c	1.6	(0.3)	25.1	(1.1)	69.4	(1.1)	3.8	(0.4)	0.1	(0.1)
Germany	0.3	(0.1)	8.2	(0.6)	50.2	(1.0)	40.4	(1.1)	0.8	(0.4)	0.0	c
Greece	0.3	(0.1)	0.5	(0.1)	3.1	(0.7)	96.1	(0.8)	0.0	c	0.0	c
Hungary	1.8	(0.7)	5.7	(0.8)	68.4	(1.1)	24.1	(0.8)	0.0	c	0.0	c
Iceland	0.0	c	0.0	c	0.0	c	100.0	c	0.0	c	0.0	c
Ireland	0.1	(0.1)	1.4	(0.2)	57.3	(1.0)	27.6	(1.4)	13.7	(1.2)	0.0	c
Israel	0.0	(0.0)	0.2	(0.1)	15.5	(1.0)	83.8	(1.0)	0.4	(0.1)	0.0	c
Italy	0.3	(0.1)	1.2	(0.2)	14.0	(0.6)	81.5	(0.8)	3.0	(0.3)	0.0	(0.0)
Japan	0.0	c	0.0	c	0.0	c	100.0	c	0.0	c	0.0	c
Korea	0.0	c	0.0	c	5.4	(1.1)	94.4	(1.1)	0.2	(0.1)	0.0	c
Luxembourg	0.7	(0.1)	9.7	(0.2)	50.2	(0.2)	39.0	(0.2)	0.4	(0.1)	0.0	c
Mexico	0.8	(0.1)	4.1	(0.3)	28.7	(1.0)	64.2	(1.1)	2.1	(0.3)	0.1	(0.1)
Netherlands	0.0	c	2.7	(0.4)	43.8	(1.1)	53.0	(1.1)	0.5	(0.2)	0.0	c
New Zealand	0.0	c	0.0	c	0.1	(0.1)	5.3	(0.4)	88.6	(0.6)	5.9	(0.6)
Norway	0.0	c	0.0	c	0.2	(0.1)	99.8	(0.1)	0.0	c	0.0	c
Poland	0.2	(0.1)	2.6	(0.3)	96.7	(0.4)	0.6	(0.2)	0.0	c	0.0	c
Portugal	2.2	(0.3)	6.6	(0.7)	27.2	(1.6)	63.8	(2.2)	0.2	(0.1)	0.0	c
Slovak Republic	1.9	(0.5)	3.5	(0.5)	38.8	(1.9)	54.0	(1.9)	1.8	(0.5)	0.0	c
Slovenia	0.0	c	0.2	(0.2)	3.8	(0.9)	91.2	(1.0)	4.7	(0.5)	0.0	c
Spain	0.1	(0.0)	7.8	(0.5)	22.3	(0.7)	69.9	(0.8)	0.0	(0.0)	0.0	c
Sweden	0.0	c	2.8	(0.3)	94.4	(0.6)	2.8	(0.6)	0.0	c	0.0	c
Switzerland	0.6	(0.2)	11.9	(1.0)	60.7	(1.7)	26.6	(1.8)	0.2	(0.1)	0.0	c
Turkey	0.7	(0.3)	1.7	(0.3)	21.9	(1.2)	70.8	(1.1)	4.8	(0.4)	0.2	(0.1)
United Kingdom	0.0	c	0.0	c	0.0	(0.0)	1.0	(0.3)	95.4	(0.3)	3.6	(0.2)
United States	0.0	c	0.1	(0.1)	8.8	(1.2)	72.7	(1.3)	18.3	(0.9)	0.2	(0.1)
OECD average	0.4	(0.0)	3.9	(0.1)	33.7	(0.2)	53.8	(0.2)	7.9	(0.1)	0.3	(0.1)
Partners												
Albania	0.1	(0.1)	1.4	(0.4)	35.7	(2.6)	62.5	(2.6)	0.3	(0.1)	0.0	c
Argentina	1.2	(0.3)	9.1	(0.9)	19.7	(1.3)	65.8	(1.9)	2.7	(0.4)	1.4	(0.8)
Brazil	0.0	c	5.0	(0.4)	11.5	(0.7)	33.8	(1.0)	46.4	(1.1)	3.3	(0.2)
Bulgaria	0.5	(0.2)	3.3	(0.5)	90.9	(0.7)	5.2	(0.5)	0.0	(0.0)	0.0	c
Colombia	3.9	(0.6)	10.8	(0.7)	21.0	(0.9)	41.4	(1.1)	22.9	(1.1)	0.0	c
Costa Rica	5.7	(0.8)	11.3	(0.8)	40.5	(1.3)	42.1	(1.7)	0.4	(0.2)	0.0	c
Croatia	0.0	c	0.0	c	77.5	(0.6)	22.5	(0.6)	0.0	c	0.0	c
Cyprus*	0.0	c	0.5	(0.1)	4.2	(0.2)	94.6	(0.2)	0.7	(0.1)	0.0	(0.0)
Hong Kong-China	0.9	(0.2)	6.0	(0.6)	24.2	(0.8)	67.3	(1.0)	1.6	(1.5)	0.0	c
Indonesia	1.5	(0.4)	6.4	(0.8)	36.8	(2.9)	50.0	(3.0)	4.7	(0.8)	0.5	(0.5)
Jordan	0.0	(0.0)	1.3	(0.2)	6.3	(0.5)	92.4	(0.6)	0.0	c	0.0	c
Kazakhstan	0.1	(0.1)	4.4	(0.5)	65.9	(1.9)	29.3	(2.1)	0.2	(0.1)	0.0	c
Latvia	0.6	(0.2)	11.6	(0.8)	83.7	(1.1)	4.1	(0.7)	0.0	c	0.0	c
Liechtenstein	5.3	(1.3)	11.5	(1.9)	62.8	(1.9)	20.4	(0.8)	0.0	c	0.0	c
Lithuania	0.1	(0.1)	5.2	(0.6)	80.2	(0.9)	14.4	(0.8)	0.0	(0.0)	0.0	c
Macao-China	3.5	(0.1)	13.3	(0.2)	33.1	(0.3)	49.5	(0.3)	0.7	(0.2)	0.0	c
Malaysia	0.0	c	0.0	c	2.9	(0.4)	97.1	(0.4)	0.0	(0.1)	0.0	c
Montenegro	0.0	c	0.0	c	77.1	(0.3)	22.9	(0.3)	0.0	c	0.0	c
Peru	2.3	(0.5)	6.6	(0.6)	16.8	(1.0)	49.1	(1.2)	25.3	(1.0)	0.0	c
Qatar	0.5	(0.1)	2.7	(0.1)	13.6	(0.1)	64.9	(0.2)	18.2	(0.1)	0.2	(0.0)
Romania	0.1	(0.1)	8.3	(0.6)	85.9	(0.9)	5.7	(0.6)	0.0	c	0.0	c
Russian Federation	0.6	(0.2)	7.3	(0.5)	73.9	(2.0)	18.1	(2.0)	0.1	(0.1)	0.0	c
Serbia	0.1	(0.1)	1.0	(0.6)	96.8	(0.7)	2.0	(0.3)	0.0	c	0.0	c
Shanghai-China	0.8	(0.2)	3.8	(0.5)	37.6	(1.8)	57.0	(1.8)	0.6	(0.1)	0.1	(0.1)
Singapore	0.4	(0.1)	2.1	(0.2)	7.6	(0.4)	89.8	(0.4)	0.2	(0.1)	0.0	c
Chinese Taipei	0.0	c	0.1	(0.1)	35.0	(1.5)	64.9	(1.4)	0.0	c	0.0	c
Thailand	0.0	(0.0)	0.2	(0.1)	19.0	(1.2)	77.5	(1.2)	3.3	(0.5)	0.0	c
Tunisia	3.9	(0.5)	9.3	(1.1)	19.4	(1.5)	60.6	(2.5)	6.7	(0.6)	0.0	c
United Arab Emirates	0.6	(0.1)	2.6	(0.4)	9.7	(1.1)	63.4	(1.7)	22.6	(1.3)	1.2	(0.3)
Uruguay	4.6	(0.6)	11.4	(0.8)	21.0	(1.1)	61.7	(1.5)	1.4	(0.2)	0.0	c
Viet Nam	0.1	(0.1)	2.1	(0.6)	6.4	(1.5)	91.4	(1.9)	0.0	c	0.0	c

Information for the adjudicated regions is available on line.

* See notes at the beginning of this Annex.

StatLink  <http://dx.doi.org/10.1787/888932937092>



ANNEX A3

TECHNICAL NOTES ON ANALYSES IN THIS VOLUME

Methods and definitions

Relative risk or increased likelihood

The relative risk is a measure of the association between an antecedent factor and an outcome factor. The relative risk is simply the ratio of two risks, i.e. the risk of observing the outcome when the antecedent is present and the risk of observing the outcome when the antecedent is not present. Figure A3.1 presents the notation that is used in the following.

■ Figure A3.1 ■

Labels used in a two-way table

P_{11}	P_{12}	$P_{1.}$
P_{21}	P_{22}	$P_{2.}$
$P_{.1}$	$P_{.2}$	$P_{..}$

$p_{..}$ is equal to $\frac{n_{..}}{n_{..}}$, with $n_{..}$ the total number of students and $p_{..}$ is therefore equal to 1, $P_{i.}$, $P_{.j}$ respectively represent the marginal probabilities for each row and for each column. The marginal probabilities are equal to the marginal frequencies divided by the total number of students. Finally, the P_{ij} represents the probabilities for each cell and are equal to the number of observations in a particular cell divided by the total number of observations.

In PISA, the rows represent the antecedent factor, with the first row for “having the antecedent” and the second row for “not having the antecedent”. The columns represent the outcome: the first column for “having the outcome” and the second column for “not having the outcome”. The relative risk is then equal to:

$$RR = \frac{(P_{11}/P_{1.})}{(P_{21}/P_{2.})}$$

Attributable risk or population relevance

The attributable risk, also referred to as population relevance in the text and tables of this volume, is interpreted as follows: if the risk factor could be eliminated, then the rate of occurrence of the outcome characteristic in the population would be reduced by this coefficient. The attributable risk is equal to (see Figure A3.1 for the notation that is used in the following formula):

$$AR = \frac{(P_{11}P_{22}) - (P_{12}P_{21})}{(P_{.1}P_{.2})}$$

The coefficients are multiplied by 100 to express the result as a percentage.

Statistics based on multilevel models

Statistics based on multi level models include variance components (between- and within-school variance), the index of inclusion derived from these components, and regression coefficients where this has been indicated. Multilevel models are generally specified as two-level regression models (the student and school levels), with normally distributed residuals, and estimated with maximum likelihood estimation. Where the dependent variable is mathematics performance, the estimation uses five plausible values for each student’s performance on the mathematics scale. Models were estimated using Mplus® software.

In multilevel models, weights are used at both the student and school levels. The purpose of these weights is to account for differences in the probabilities of students being selected in the sample. Since PISA applies a two-stage sampling procedure, these differences are due to factors at both the school and the student levels. For the multilevel models, student final weights (W_FSTUWT) were used. Within-school-weights correspond to student final weights, rescaled to sum up within each school to the school sample size. Between-school weights correspond to the sum of student final weights (W_FSTUWT) within each school. The definition of between-school weights has changed with respect to PISA 2009.

The index of inclusion is defined and estimated as:

$$100 * \frac{\sigma_w^2}{\sigma_w^2 + \sigma_b^2}$$

where σ_w^2 and σ_b^2 , respectively, represent the within- and between-variance estimates.



The results in multilevel models, and the between-school variance estimate in particular, depend on how schools are defined and organised within countries and by the units that were chosen for sampling purposes. For example, in some countries, some of the schools in the PISA sample were defined as administrative units (even if they spanned several geographically separate institutions, as in Italy); in others they were defined as those parts of larger educational institutions that serve 15-year-olds; in still others they were defined as physical school buildings; and in others they were defined from a management perspective (e.g. entities having a principal). The *PISA 2012 Technical Report* (OECD, forthcoming) and Annex A2 provide an overview of how schools were defined. In Slovenia, the primary sampling unit is defined as a group of students who follow the same study programme within a school (an educational track within a school). So in this particular case the between-school variance is actually the within-school, between-track variation. The use of stratification variables in the selection of schools may also affect the estimate of the between-school variance, particularly if stratification variables are associated with between-school differences.

Because of the manner in which students were sampled, the within-school variation includes variation between classes as well as between students.

Multiple imputation replaces each missing value with a set of plausible values that represent the uncertainty about the right value to impute. The multiple imputed data sets are then analysed by using standard procedures for complete data and by combining results from these analyses. Five imputed values are computed for each missing value. Different methods can be used according to the pattern of missing values. For arbitrary missing data patterns, the MCMC (Monte Carlo Markov Chain) approach can be used.

This approach is used with the SAS procedure MI for the multilevel analyses in this volume. Multiple imputation is conducted separately for each model and each country, except for the model with all variables (Tables IV.1.12a, IV.1.12b and IV.1.12c) in which the data were constructed from imputed data for the individual models, such as the model for learning environment, model for selecting and grouping students, etc. Where continuous values are generated for missing discrete variables, these are rounded to the nearest discrete value of the variable. Each of the five plausible value of mathematics performance is analysed by Mplus[®] software using one of the five imputed data sets, which were combined taking account of the between imputation variance.

Standard errors and significance tests

The statistics in this report represent estimates of national performance based on samples of students, rather than values that could be calculated if every student in every country had answered every question. Consequently, it is important to measure the degree of uncertainty of the estimates. In PISA, each estimate has an associated degree of uncertainty, which is expressed through a standard error. The use of confidence intervals provides a way to make inferences about the population means and proportions in a manner that reflects the uncertainty associated with the sample estimates. From an observed sample statistic and assuming a normal distribution, it can be inferred that the corresponding population result would lie within the confidence interval in 95 out of 100 replications of the measurement on different samples drawn from the same population.

In many cases, readers are primarily interested in whether a given value in a particular country is different from a second value in the same or another country, e.g. whether girls in a country perform better than boys in the same country. In the tables and charts used in this report, differences are labelled as statistically significant when a difference of that size, smaller or larger, would be observed less than 5% of the time, if there were actually no difference in corresponding population values. Similarly, the risk of reporting a correlation as significant if there is, in fact, no correlation between two measures, is contained at 5%.

Throughout the report, significance tests were undertaken to assess the statistical significance of the comparisons made.

Gender differences and differences between subgroup means

Gender differences in student performance or other indices were tested for statistical significance. Positive differences indicate higher scores for boys while negative differences indicate higher scores for girls. Generally, differences marked in bold in the tables in this volume are statistically significant at the 95% confidence level.

Similarly, differences between other groups of students (e.g. native students and students with an immigrant background) were tested for statistical significance. The definitions of the subgroups can in general be found in the tables and the text accompanying the analysis. All differences marked in bold in the tables presented in Annex B of this report are statistically significant at the 95% level.

Differences between subgroup means, after accounting for other variables

For many tables, subgroup comparisons were performed both on the observed difference (“before accounting for other variables”) and after accounting for other variables, such as the *PISA index of economic, social and cultural status of students* (ESCS). The adjusted differences were estimated using linear regression and tested for significance at the 95% confidence level. Significant differences are marked in bold.

Performance differences between the top and bottom quartiles of PISA indices and scales

Differences in average performance between the top and bottom quarters of the PISA indices and scales were tested for statistical significance. Figures marked in bold indicate that performance between the top and bottom quarters of students on the respective index is statistically significantly different at the 95% confidence level.



Differences between subgroups of schools

In this Volume, schools are compared across several aspects, such as resource allocation or performance. For this purpose, schools are grouped in categories by socio-economic status of students and schools, public-private status, lower and upper secondary education and school location. The differences between subgroups of schools are tested for statistical significance in the following way:

- *Socio-economic status of students*: Students in the top quarter of ESCS are compared to students in the bottom quarter of ESCS. If the difference is statistically significant at the 95% confidence levels, both figures are marked in bold. The second and third quarters do not enter the comparison.
- *Socio-economic status of schools*: advantaged schools are compared to disadvantaged schools. If the difference is statistically significant at the 95% confidence levels, both figures are marked in bold. Average schools do not enter the comparison.
- *Public and private schools*: Government-dependent and government-independent private schools are jointly considered as private schools. Figures in bold in data tables presented in Annex B of this report indicate statistically significant differences, at the 95% confidence level, between public and private schools.
- *Education levels*: Students at the upper secondary education are compared to students at the lower secondary education. If the difference is statistically significant at the 95% confidence levels, both figures are marked in bold.
- *School location*: For the purpose of significance tests, “schools located in a small town” and “schools located in a town” are jointly considered to form a single group. Figures for “schools located in a city or large city” are marked in bold in data tables presented in Annex B of this report if the difference with this middle category (“schools located in a small town” and “schools located in a town”) is significant at the 95% confidence levels. In turn, figures for “schools located in a village, hamlet, or rural area” are marked in bold if the difference with this middle category is significant. Differences between the extreme categories were not tested for significance.

Change in the performance per unit of the index

For many tables, the difference in student performance per unit of the index shown was calculated. Figures in bold indicate that the differences are statistically significantly different from zero at the 95% confidence level.

Relative risk or increased likelihood

Figures in bold in the data tables presented in Annex B of this report indicate that the relative risk is statistically significantly different from 1 at the 95% confidence level. To compute statistical significance around the value of 1 (the null hypothesis), the relative-risk statistic is assumed to follow a log-normal distribution, rather than a normal distribution, under the null hypothesis.

Attributable risk or population relevance

Figures in bold in the data tables presented in Annex B of this report indicate that the attributable risk is statistically significantly different from 0 at the 95% confidence level.

Standard errors in statistics estimated from multilevel models

For statistics based on multilevel models (such as the estimates of variance components and regression coefficients from two-level regression models) the standard errors are not estimated with the usual replication method which accounts for stratification and sampling rates from finite populations. Instead, standard errors are “model-based”: their computation assumes that schools, and students within schools, are sampled at random (with sampling probabilities reflected in school and student weights) from a theoretical, infinite population of schools and students which complies with the model’s parametric assumptions.

The standard error for the estimated index of inclusion is calculated by deriving an approximate distribution for it from the (model-based) standard errors for the variance components, using the delta-method.

Standard errors in trend analyses of performance: Link error

Standard errors for performance trend estimates had to be adjusted because the equating procedure that allows scores in different PISA assessments to be compared introduces a form of random error that is related to performance changes on the link items. These more conservative standard errors (larger than standard errors that were estimated before the introduction of the link error) reflect not only the measurement precision and sampling variation as for the usual PISA results, but also the link error (see Annex A5 for a technical discussion of the link error).

Link items represent only a subset of all items used to derive PISA scores. If different items were chosen to equate PISA scores over time, the comparison of performance for a group of students across time could vary. As a result, standard errors for the estimates of the change over time in mathematics, reading or science performance of a particular group (e.g. a country or economy, a region, boys, girls, students with an immigrant background, students without an immigrant background, socio-economically advantaged students, students in public schools, etc.) include the link error in addition to the sampling and imputation error commonly added to estimates in performance for a particular year. Because the equating procedure adds uncertainty to the position in the distribution (a change in the intercept) but does not result in any change in the variance of a distribution, standard errors for location-invariant estimates do not



include the link error. Location-invariant estimates include, for example, estimates for variances, regression coefficients for student- or school-level covariates, and correlation coefficients.

Figures in bold in the data tables for trends in performance presented in Annex B of this report indicate that the the change in performance for that particular group is statistically significantly different from 0 at the 95% confidence level. The standard errors used to calculate the statistical significance of the reported trend include the link error.



ANNEX A4

QUALITY ASSURANCE

Quality assurance procedures were implemented in all parts of PISA 2012, as was done for all previous PISA surveys.

The consistent quality and linguistic equivalence of the PISA 2012 assessment instruments were facilitated by providing countries with equivalent source versions of the assessment instruments in English and French and requiring countries (other than those assessing students in English and French) to prepare and consolidate two independent translations using both source versions. Precise translation and adaptation guidelines were supplied, also including instructions for selecting and training the translators. For each country, the translation and format of the assessment instruments (including test materials, marking guides, questionnaires and manuals) were verified by expert translators appointed by the PISA Consortium before they were used in the PISA 2012 Field Trial and Main Study. These translators' mother tongue was the language of instruction in the country concerned and they were knowledgeable about education systems. For further information on the PISA translation procedures, see the *PISA 2012 Technical Report* (OECD, forthcoming).

The survey was implemented through standardised procedures. The PISA Consortium provided comprehensive manuals that explained the implementation of the survey, including precise instructions for the work of School Co-ordinators and scripts for Test Administrators to use during the assessment sessions. Proposed adaptations to survey procedures, or proposed modifications to the assessment session script, were submitted to the PISA Consortium for approval prior to verification. The PISA Consortium then verified the national translation and adaptation of these manuals.

To establish the credibility of PISA as valid and unbiased and to encourage uniformity in administering the assessment sessions, Test Administrators in participating countries were selected using the following criteria: it was required that the Test Administrator not be the reading, mathematics or science instructor of any students in the sessions he or she would administer for PISA; it was recommended that the Test Administrator not be a member of the staff of any school where he or she would administer for PISA; and it was considered preferable that the Test Administrator not be a member of the staff of any school in the PISA sample. Participating countries organised an in-person training session for Test Administrators.

Participating countries and economies were required to ensure that: Test Administrators worked with the School Co-ordinator to prepare the assessment session, including updating student tracking forms and identifying excluded students; no extra time was given for the cognitive items (while it was permissible to give extra time for the student questionnaire); no instrument was administered before the two one-hour parts of the cognitive session; Test Administrators recorded the student participation status on the student tracking forms and filled in a Session Report Form; no cognitive instrument was permitted to be photocopied; no cognitive instrument could be viewed by school staff before the assessment session; and Test Administrators returned the material to the national centre immediately after the assessment sessions.

National Project Managers were encouraged to organise a follow-up session when more than 15% of the PISA sample was not able to attend the original assessment session.

National Quality Monitors from the PISA Consortium visited all national centres to review data-collection procedures. Finally, School Quality Monitors from the PISA Consortium visited a sample of seven schools during the assessment. For further information on the field operations, see the *PISA 2012 Technical Report* (OECD, forthcoming).

Marking procedures were designed to ensure consistent and accurate application of the marking guides outlined in the PISA Operations Manuals. National Project Managers were required to submit proposed modifications to these procedures to the Consortium for approval. Reliability studies to analyse the consistency of marking were implemented.

Software specially designed for PISA facilitated data entry, detected common errors during data entry, and facilitated the process of data cleaning. Training sessions familiarised National Project Managers with these procedures.

For a description of the quality assurance procedures applied in PISA and in the results, see the *PISA 2012 Technical Report* (OECD, forthcoming).

The results of adjudication showed that the PISA Technical Standards were fully met in all countries and economies that participated in PISA 2012, with the exception of Albania. Albania submitted parental occupation data that was incomplete and appeared inaccurate, since there was over-use of a narrow range of occupations. It was not possible to resolve these issues during the course of data cleaning, and as a result neither parental occupation data nor any indices which depend on this data are included in the international dataset. Results for Albania are omitted from any analyses which depend on these indices.

ANNEX A5

TECHNICAL DETAILS OF TRENDS ANALYSES

Comparing mathematics, reading and science performance across PISA cycles

The PISA 2003, 2006, 2009 and 2012 assessments use the same mathematics performance scale, which means that score points on this scale are directly comparable over time. The same is true for the reading performance scale used since PISA 2000 and the science performance scale used since PISA 2006. The comparability of scores across time is possible because of the use of link items that are common across assessments and can be used in the equating procedure to align performance scales. The items that are common across assessments are a subset of the total items that make up the assessment because PISA progressively renews its pool of items. As a result, out of a total of 110 items in the PISA 2012 mathematics assessment, 84 are linked to 2003 items, 48 to 2006 items and 35 to 2009 items. The number of PISA 2012 items linked to the PISA 2003 assessment is larger than the number linked to the PISA 2006 or the PISA 2009 assessments because mathematics was a major domain in PISA 2003 and PISA 2012. In PISA 2006 and PISA 2009, mathematics was a minor domain and all the mathematics items included in these assessments were link items. The *PISA 2012 Technical Report* (OECD, forthcoming) provides the technical details on equating the PISA 2012 mathematics scale for trends purposes.

Link error

Standard errors for performance trend estimates had to be adjusted because the equating procedure that allows scores in different PISA assessments to be compared introduces a form of random error that is related to performance changes on the link items. These more conservative standard errors (larger than standard errors that were estimated before the introduction of the link error) reflect not only the measurement precision and sampling variation as for the usual PISA results, but also the link error provided in Table A5.1.

Link items represent only a subset of all items used to derive PISA scores. If different items were chosen to equate PISA scores over time, the comparison of performance for a group of students across time could vary. As a result, standard errors for the estimates of the change over time in mathematics, reading or science performance of a particular group (e.g. a country or economy, a region, boys, girls, students with an immigrant background, students without an immigrant background, socio-economically advantaged students, students in public schools, etc.) include the link error in addition to the sampling and imputation error commonly added to estimates in performance for a particular year. Because the equating procedure adds uncertainty to the position in the distribution (a change in the intercept) but does not result in any change in the variance of a distribution, standard errors for location-invariant estimates do not include the link error. Location-invariant estimates include, for example, estimates for variances, regression coefficients for student- or school-level covariates, and correlation coefficients.

Link error for scores between two PISA assessments

The following equations describe how link errors between two PISA assessments are calculated. Suppose we have L score points in K units. Use i to index items in a unit and j to index units so that $\hat{\mu}_{ij}^y$ is the estimated difficulty of item i in unit j for year y , and let for example to compare PISA 2006 and PISA 2003:

$$c_{ij} = \hat{\mu}_{ij}^{2006} - \hat{\mu}_{ij}^{2003}$$

The size (total number of score points) of unit j is m_j so that:

$$\sum_{j=1}^K m_j = L$$

and

$$\bar{m} = \frac{1}{K} \sum_{j=1}^K m_j$$

Further let:

$$c_j = \frac{1}{m_j} \sum_{i=1}^{m_j} c_{ij}$$

and

$$\bar{c} = \frac{1}{N} \sum_{j=1}^K \sum_{i=1}^{m_j} c_{ij}$$

then the link error, taking clustering into account, is as follows:

$$error_{2006,2003} = \sqrt{\frac{\sum_{j=1}^K m_j^2 (c_j - \bar{c})^2}{K(K-1)\bar{m}^2}}$$


This approach for estimating the link errors was used in PISA 2006, PISA 2009 and PISA 2012. The link errors for comparisons of PISA 2012 results with previous assessments are shown in Table A5.1.



[Part 1/1]
Table A5.1 Link error for comparisons of performance between PISA 2012 and previous assessments

Comparison	Mathematics	Reading	Science
PISA 2000 to PISA 2012		5.923	
PISA 2003 to PISA 2012	1.931	5.604	
PISA 2006 to PISA 2012	2.084	5.580	3.512
PISA 2009 to PISA 2012	2.294	2.602	2.006

Note: Comparisons between PISA 2012 scores and previous assessments can only be made to when the subject first became a major domain. As a result, comparisons in mathematics performance between PISA 2012 and PISA 2000 are not possible, nor are comparisons in science performance between PISA 2012 and PISA 2000 or PISA 2003.

StatLink  <http://dx.doi.org/10.1787/888932960500>

Comparisons of performance: Difference between two assessments

To evaluate the evolution of performance, analyses report the change in performance between two cycles. Comparisons between two assessments (e.g. a country's/economy's change in performance between PISA 2003 and PISA 2012 or the change in performance of a subgroup) are calculated as:

$$\Delta_{2012-t} = PISA_{2012} - PISA_t$$

where Δ_{2012-t} is the difference in performance between PISA 2012 and a previous PISA assessment, where t can take any of the following values: 2000, 2003, 2006 or 2009. $PISA_{2012}$ is the mathematics, reading or science score observed in PISA 2012, and $PISA_t$ is the mathematics, reading or science score observed in a previous assessment (2000, 2003, 2006 or 2009). The standard error of the change in performance $\sigma(\Delta_{2012-t})$ is:

$$\sigma(\Delta_{2012-t}) = \sqrt{\sigma_{2012}^2 + \sigma_t^2 + error_{2012,t}^2}$$

where σ_{2012} is the standard error observed for $PISA_{2012}$, σ_t is the standard error observed for $PISA_t$ and $error_{2012,t}$ is the link error for comparisons of mathematics, reading or science performance between the PISA 2012 assessment and a previous (t) assessment. The value for $error_{2012,t}$ is shown in Table A5.1.

Comparing items and non-performance scales across PISA cycles

To gather information about students' and schools' characteristics, PISA asks both students and schools to complete a background questionnaire. In PISA 2003 and PISA 2012 several questions were left untouched, allowing for a comparison of responses to these questions over time. In this report, only questions that retained the same wording were used for trends analyses. Questions with subtle word changes or questions with major word changes were not compared across time because it is impossible to discern whether observed changes in the response are due to changes in the construct they are measuring or to changes in the way the construct is being measured.

Also, as described in Annex A1, questionnaire items in PISA are used to construct indices. Whenever the questions used in the construction of indices remains intact in PISA 2003 and PISA 2012, the corresponding indices are compared. Two types of indices are used in PISA: simple indices and scale indices.

Simple indices recode a set of responses to questionnaire items. For trends analyses, the values observed in PISA 2003 are compared directly to PISA 2012, just as simple responses to questionnaire items are. This is the case of indices like student-teacher ratio and ability grouping in mathematics.

Scale indices, on the other hand, imply WLE estimates which require rescaling in order to be comparable across PISA cycles. Scale indices, like the *PISA index of economic, social and cultural status*, the *index of sense of belonging*, the *index of attitudes towards school*, the *index of intrinsic motivation to learn mathematics*, the *index of instrumental motivation to learn mathematics*, the *index of mathematics self-efficacy*, the *index of mathematics self-concept*, the *index of anxiety towards mathematics*, the *index of teacher shortage*, the *index of quality of physical infrastructure*, the *index of quality of educational resources*, the *index of disciplinary climate*, the *index of student-teacher relations*, the *index of teacher morale*, the *index of student-related factors affecting school climate*, and the *index of teacher-related factors affecting school climate*, were scaled in PISA 2012 to have an OECD mean of 0 and a standard deviation of 1. In PISA 2003 these same scales were scaled to have an OECD average of 0 and a standard deviation of 1. Because they are on different scales, values reported in *Learning for Tomorrow's World: First Results from PISA 2003* (OECD, 2004) cannot be compared with those reported in this volume. To make these scale indices comparable, values for 2003 have been rescaled to the 2012 scale, using the PISA 2012 parameter estimates.



To evaluate change in these items and scales, analyses report the change in the estimate between two assessments, usually PISA 2003 and PISA 2012. Comparisons between two assessments (e.g. a country's/economy's change index of anxiety towards mathematics between PISA 2003 and PISA 2012 or the change in this index for a subgroup) is calculated as:

$$\Delta_{2012,t} = PISA_{2012} - PISA_t$$

where $\Delta_{2012,t}$ is the difference in the index between PISA 2012 and a previous assessment, $PISA_{2012}$ is the index value observed in PISA 2012, and $PISA_t$ is the index value observed in a previous assessment (2000, 2003, 2006 or 2009). The standard error of the change in performance $\sigma(\Delta_{2012-t})$ is:

$$\sigma(\Delta_{2012-t}) = \sqrt{\sigma_{2012}^2 + \sigma_t^2}$$

where σ_{2012} is the standard error observed for $PISA_{2012}$ and σ_t is the standard error observed for $PISA_t$. These comparisons are based on an identical set of items; there is no uncertainty related to the choice of items for equating purposes, so no link error is needed.

Although only scale indices that use the same items in PISA 2003 and PISA 2012 are valid for trend comparisons, this does not imply that PISA 2012 indices that include exactly the same items as 2003 as well as new questionnaire items cannot be compared with PISA 2003 indices that included a smaller pool of items. In such cases, for example the *index of sense of belonging*, trend analyses were conducted by treating as missing in PISA 2003 items that were asked in the context of PISA 2012 but not in the PISA 2003 student questionnaire. This means that while the full set of information was used to scale the sense of belonging index in 2012, the PISA 2003 sense of belonging index was scaled under the assumption that if the 2012 items that were missing in 2003 had been asked in 2003, the overall index and index variation would have remained the same as those that were observed on common 2003 items. This is a tenable assumption inasmuch as in both PISA 2003 and PISA 2012 the questionnaire items used to construct the scale hold as an underlying factor in the construction of the scale.

OECD average

Throughout this report, the OECD average is used as a benchmark. It is calculated as the average across OECD countries, weighting each country equally. Some OECD countries did not participate in certain assessments, other OECD countries do not have comparable results for some assessments, others did not include certain questions in their questionnaires or changed them substantially from assessment to assessment. For this reason in trends tables and figures, the OECD average is reported as assessment-specific, that is, it includes only those countries for which there is comparable information in that particular assessment. This way, the 2003 OECD average includes only those OECD countries that have comparable information from the 2003 assessment, even if the results it refers to the PISA 2012 assessment and more countries have comparable information. This restriction allows for valid comparisons of the OECD average over time.

References

OECD (forthcoming), *PISA 2012 Technical Report, PISA*, OECD Publishing.

OECD (2004), *Learning for Tomorrow's World: First Results from PISA 2003*, PISA, OECD Publishing.
<http://dx.doi.org/10.1787/9789264006416-en>



ANNEX A6

ANCHORING VIGNETTES IN THE PISA 2012 STUDENT QUESTIONNAIRE

Annex A6 is available on line only.

It can be found at: www.pisa.oecd.org