Annex A

PISA 2015 TECHNICAL BACKGROUND

All tables in Annex A are available on line

Annex A1: Indices from the student questionnaire

Annex A2: The PISA target population, the PISA samples and the definition of schools
http://dx.doi.org/10.1787/888933433129

Annex A3: Technical notes on analyses in this volume

Annex A4: Quality assurance

Annex A5: Changes in the administration and scaling of PISA 2015 and implications for trends analyses

Annex A6: Guidelines and caveats about interpreting the results

Note regarding B-S-J-G (China)
B-S-J-G (China) refers to the four PISA participating China provinces : Beijing, Shanghai, Jiangsu, Guangdong.

Note regarding CABA (Argentina)
CABA (Argentina) refers to the Ciudad Autónoma de Buenos Aires, Argentina.

Note regarding FYROM
FYROM refers to the Former Yugoslav Republic of Macedonia.

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
INDICES FROM THE STUDENT QUESTIONNAIRE

Explanation of the indices

This section explains the indices derived from the PISA 2015 student and school context questionnaires used in this volume.

Several PISA measures reflect indices that summarise responses from students, their parents, teachers 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 2015 Assessment and Analytical Framework* (OECD, 2016) 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. 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 the *PISA 2015 Technical Report* (OECD, forthcoming).

There are three types of indices: simple indices, new scale indices, and trend 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.

New and trend scale indices are the variables constructed through the scaling of multiple items. Unless otherwise indicated, the index was scaled using a two-parameter item response model (a generalised partial credit model was used in the case of items with more than two categories) and values of the index correspond to Warm likelihood estimates (WLE) (Warm, 1989). For details on how each scale index was constructed, see the *PISA 2015 Technical Report* (OECD, forthcoming). In general, the scaling was done in three stages:

1. The item parameters were estimated from equally-weighted samples of students from all countries and economies; only cases with a minimum number of three valid responses to items that are part of the index were included. In the case of **trend indices**, a common calibration linking procedure was used: countries/economies that participated in both PISA 2006 and PISA 2015 contributed both samples to the calibration of item parameters; each cycle, and, within each cycle, each country/economy contributed equally to the estimation.
2. The estimates were computed for all students and all schools by anchoring the item parameters obtained in the preceding step.
3. For **new scale indices**, the Warm likelihood estimates 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). **Trend indices** were equated so that the mean and standard deviation across OECD countries of rescaled PISA 2006 estimates and of the original estimates included in the PISA 2006 database matched. Trend indices are therefore reported on the same scale as used originally in PISA 2006, so that values can be directly compared to those included in the PISA 2006 database.

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, on average, than respondents in OECD countries did. 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 were used in this volume and 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 and "SC" for the items in the school questionnaire. All the context questionnaires, and the PISA international database, including all variables, are available through www.oecd.org/pisa.

Student-level simple indices

Student age

The age of a student (AGE) was calculated as the difference between the year and month of the testing and the year and month of a student's birth. Data on student's age were obtained from both the questionnaire (ST003) and the student tracking forms. If the month of testing was not known for a particular student, the median month for that country was used in the calculation.

Parents' level of education

Students' responses on questions ST005, ST006, ST007 and ST008 regarding parental education were classified using ISCED 1997 (OECD, 1999). Indices on parental education were constructed by recoding educational qualifications into the following categories: (0) None, (1) <ISCED level 1> (primary education), (2) <ISCED level 2> (lower secondary), (3) <ISCED level 3B or 3C> (vocational/pre-vocational upper secondary), (4) <ISCED level 3A> (general upper secondary) and/or <ISCED level 4> (non-tertiary post-secondary), (5) <ISCED level 5B> (vocational tertiary) and (6) <ISCED level 5A> and/or <ISCED level 6> (theoretically oriented tertiary and post-graduate). Indices with these categories were provided for a student's mother (MISCED) and father (FISCED). In addition, the index of highest education level of parents (HISCED) corresponds to the higher ISCED level of either parent. The index of highest education level of parents was also recoded into estimated number of years of schooling (PARED). The correspondence between education levels and years of schooling is available in the *PISA 2015 Technical Report* (OECD, forthcoming).

Parents' highest occupational status

Occupational data for both the student's father and the student's mother were obtained from responses to open-ended questions. The responses were coded to four-digit ISCO codes (ILO, 2007) and then mapped to the international socio-economic index of occupational status (ISEI) (Ganzeboom and Treiman, 2003). In PISA 2015, as in PISA 2012, the new ISCO and ISEI in their 2008 version were used rather than the 1988 versions that had been applied in the previous four cycles (Ganzeboom, 2010). Three indices were calculated based on this information: father's occupational status (BFMJ2); mother's occupational status (BMMJ1); and the highest occupational status of parents (HISEI) which corresponds to the higher ISEI score of either parent or to the only available parent's ISEI score. For all three indices, higher ISEI scores indicate higher levels of occupational status.

Immigrant background

The PISA database contains three country-specific variables relating to the students' country of birth, their mother and father (COBN_S, COBN_M and COBN_F). The items ST019Q01TA, ST019Q01TB and ST019Q01TC were recoded into the following categories: (1) country of birth is the same as country of assessment and (2) other. The index of immigrant background (IMMIG) was calculated from these variables with the following categories: (1) non-immigrant students (those students who had at least one parent born in the country), (2) second-generation immigrant students (those born in the country of assessment but whose parent(s) were born in another country) and (3) first-generation immigrant students (those students 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 were assigned missing values for this variable.

Grade repetition

The grade repetition variable (REPEAT) was computed by recoding variables ST127Q01TA, ST127Q02TA and ST127Q03TA. REPEAT took the value of "1" if the student had repeated a grade in at least one ISCED level and the value of "0" if "no, never" was chosen at least once, given that none of the repeated grade categories were chosen. The index is assigned a missing value if none of the three categories were ticked in any levels.

Study programme

PISA collects data on study programmes available to 15-year old students in each country. This information is obtained through the student tracking form and the student questionnaire. In the final database, all national programmes are included in a separate derived variable (PROGN) where the first six digits represent the National Centre code, and the last two digits are the nationally specific programme code. All study programmes were classified using the International Standard Classification of Education (ISCED) (OECD, 1999). The following indices were derived from the data on study programmes:

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

Learning time

Learning time in test language regular lessons (LMINS) was computed by multiplying the number of minutes on average in the test language class by number of test language class periods per week (ST061 and ST059). Comparable indices were computed for mathematics (MMINS) and science (SMINS). Learning time in total (TMINS) was computed using information about the average minutes in a <class period> (ST061) in relation to information about the number of class periods per week attended in total (ST060). For convenience purposes, the information on learning time has been transformed into hours.

Out-of-school study time

Students were asked in a slider-format question how much time they spent studying in addition to their required school schedule (ST071). The index OUTHOURS was computed by summing the time spent studying for different school subjects.

Skipping classes or days of school

Students' 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.

Arriving late for school

Students responded to a question whether and how frequently they arrived late for school during the last two weeks before the PISA test (ST062). This variable is used to derive an indicator of student truancy which takes a value of 0 if students reported not arriving late to school or arrived to school less than 3 days in the last two weeks and takes a value of 1 if students reported arriving to school late at least three days in the same period.

Perceived teacher support

Perceived teacher support refers to students reporting "every lesson" or "most lessons" to the statements "The teacher shows an interest in every student's learning", "The teacher gives extra help when students need it" and "The teacher helps students with their learning" in their responses to a question on things that happen during their science lessons (ST100).

Perceptions of teachers behaving unfairly

Perception of teachers behaving unfairly refers to students reporting "a few times a month" or "once a week or more" to the statements "Teachers disciplined me more harshly than other students", "Teachers ridiculed me in front of others" or "Teachers said something insulting to me in front of others" in their responses to a question on their school experiences with teachers (ST039).

Science-related career expectations

In PISA 2015, students were asked to answer a question (ST114) about "what kind of job [they] expect to have when [they] are about 30 years old". Answers to this open-ended question were coded to four-digit ISCO codes (ILO, 2007), in variable OCOD3. This variable was used to derive the index of science-related career expectations.

Science-related career expectations are defined as those career expectations whose realisation requires further engagement with the study of science beyond compulsory education, typically in formal tertiary education settings. The classification of careers into science-related and non-science-related is based on the four-digit ISCO-08 classification of occupations.

Only professionals (major ISCO group 2) and technicians/associate professionals (major ISCO group 3) were considered to fit the definition of science-related career expectations. In a broad sense, several managerial occupations (major ISCO group 1) are clearly science-related: these include research and development managers, hospital managers, construction managers, and other occupations classified under production and specialised services managers (submajor group 13). However, it was considered that when science-related experience and training is an important requirement of a managerial occupation, these are not entry-level jobs and 15-year-old students with science-related career expectations would not expect to be in such a position by age 30.

Several skilled agriculture, forestry and fishery workers (major ISCO group 6) could also be considered to work in science-related occupations. The United States O*NET OnLine (2016) classification of science, technology, engineering and mathematics (STEM) occupations indeed include these occupations. These, however, do not typically require formal science-related training or study after compulsory education. On these grounds, only major occupation groups that require ISCO skill levels 3 and 4 were included among science-related occupational expectations.

Among professionals and technicians/associate professionals, the boundary between science-related and non-science related occupations is sometimes blurred, and different classifications draw different lines.

The classification used in this report includes four groups of jobs:[1]

1. *Science and engineering professionals*: All science and engineering professionals (submajor group 21), except product and garment designers (2163), graphic and multimedia designers (2166).
2. *Health professionals*: All health professionals in submajor group 22 (e.g. doctors, nurses, veterinarians), with the exception of traditional and complementary medicine professionals (minor group 223).
3. *ICT professionals*: All information and communications technology professionals (submajor group 25).

1. In the United Kingdom (excluding Scotland), career expectations were coded to the three-digit level only. As a result, the occupations of product and garment designers (ISCO08: 2163) and graphic and multimedia designers (2166) are included among science and engineering professionals, medical and dental prosthetic technicians (3214) are included among science technicians and associate professionals, while telecommunications engineering technicians (3522) are excluded. These careers represent a small percentage of the students classified as having science-related career expectations, such that results are not greatly affected.

4. *Science technicians and associate professionals*, including:
   - physical and engineering science technicians (minor group 311)
   - life science technicians and related associate professionals (minor group 314)
   - air traffic safety electronic technicians (3155)
   - medical and pharmaceutical technicians (minor group 321), except medical and dental prosthetic technicians (3214)
   - telecommunications engineering technicians (3522).

How this classification compares to existing classifications

When three existing classifications of 15-year-olds' science career expectations, all based on the International Standard Classification of Occupations (ISCO), 1988 edition (ISCO-88), are compared to the present classification, based on ISCO-08, a few differences emerge. Some are due to the updated version of occupational codings (as discussed in the next section); the remaining differences are summarised in Table A1.1.

Table A1.1 ■ **Differences in the definition of science-related career expectations**

	This classification	OECD (2007)	Sikora and Pokropek (2012)	Kjærnsli and Lie (2011)
Science-related managerial jobs	out	in	in	out
Psychologists	out	in	in	out
Sociologists and social work professionals	out	in	out	out
Photographers and image and sound recording equipment operators, broadcasting and telecommunications equipment operators	out	in	in	out
Statistical, mathematical and related associate professionals	out	out	in	out
Aircraft controllers (e.g. pilots, air traffic controllers)	out	in	in	out
Ship controllers (Ships' desk officers, etc.)	out	out	in	out
Medical assistants, dental assistants, veterinary assistants, nursing and midwifery associate professionals	out	in	in	out
Computer assistants, computer equipment operators and industrial robot controllers	out	out	out	in
Air traffic safety electronic technicians	in	in	in	out
Pharmaceutical technicians and assistants	in	in	in	out
Dieticians and nutritionists	in	in	in	out

Developing a comparable classification for ISCO-88

The same open-ended question was also included in the PISA 2006 questionnaire (ID in 2006: ST30), but students' answers were coded in the PISA 2006 database according to ISCO-88. It is not possible to ensure a strictly comparable classification. To report changes over time, the correspondence described in Table A1.2 was used to derive a similar classification based on PISA 2006 data:

Table A1.2 ■ **ISCO-08 to ISCO-88 correspondence table for science-related career expectations**

Group	ISCO-08	ISCO-88
Science and engineering professionals	21xx (except 2163 and 2166)	21xx (except 213x), 221x
Health professionals	22xx (except 223x)	22xx (except 221x), 3223, 3226
ICT professionals	25xx	213x
Science technicians and associate professionals	311x, 314x, 3155, 321x (except 3214), 3522	311x, 3133, 3145, 3151, 321x, 3228

The main differences between ISCO-88 and ISCO-08, for the purpose of deriving the index of science-related career expectations, are the following:

- Medical equipment operators (ISCO-88: 3133) correspond to medical imaging and therapeutic equipment technicians in ISCO-08; air traffic safety technicians (ISCO-88: 3145) correspond to air traffic safety electronics technicians in ISCO-08; building and fire inspectors (ISCO-88: 3151) mostly correspond to civil engineering technicians in ISCO-08.
- Dieticians and nutritionists (ISC0-88: 3223) are classified among professionals in ISCO-08. For consistency, this ISCO-88 occupation was classified among health professionals.

- Physiotherapists and related associate professionals (ISCO-88: 3226) form two distinct categories in ISCO-08, with physiotherapists classified among professionals. Given that students who expect to work as physiotherapists far outnumber those who expect to work as related associate professionals, this ISCO-88 occupation was classified among health professionals.
- Several health-related occupations classified as "modern health associate professionals" in ISCO-88 are included among health professionals in ISCO-08 (e.g. speech therapist, ophthalmic opticians). While health professionals are, in general, included among science-related careers, health associate professionals are not included among science-related careers. In applying the classification to ISCO-88, the entire code was excluded from science-related careers.
- Telecommunications engineering technicians (ISCO-08: 3522) do not form a separate occupation in ISCO-88, where they can be found among electronics and telecommunications engineering technicians (ISCO-88: 3114).
- Information and communications technology professionals form a distinct submajor group (25) in ISCO-08 but are classified among physical, mathematical and engineering science professionals in ISCO-88.

Student-level scale indices

New scale indices

Schoolwork-related anxiety

The index of schoolwork-related anxiety (ANXTEST) was constructed using student responses to question (ST118) over the extent they strongly agreed, agreed, disagreed or strongly disagreed with the following statements when asked to think about him or herself: I often worry that it will be difficult for me taking a test; I worry that I will get poor <grades> at school; Even if I am well prepared for a test I feel very anxious; I get very tense when I study; I get nervous when I don't know how to solve a task at school.

Achievement motivation

The index of achievement motivation (MOTIVAT) was constructed using students' responses to a new question developed for PISA 2015 (ST119). Students reported, on a four-point Likert scale with the answering categories "strongly disagree", "disagree", "agree", and "strongly agree", their agreement with the following statements: I want top grades in most or all of my courses; I want to be able to select from among the best opportunities available when I graduate; I want to be the best, whatever I do; I see myself as an ambitious person; I want to be one of the best students in my class. Higher values indicate that students have greater achievement motivation.

Trend scale indices

Enjoyment of science

The index of enjoyment of science (JOYSCIE) was constructed based on a trend question (ST094) from PISA 2006 (ID in 2006: ST16), asking students on a four-point Likert scale with the categories "strongly agree", "agree", "disagree", and "strongly disagree" about their agreement with the following statements: I generally have fun when I am learning <broad science> topics; I like reading about <broad science>; I am happy working on <broad science> topics; I enjoy acquiring new knowledge in <broad science>; and I am interested in learning about <broad science>. The derived variable JOYSCIE was equated to the corresponding scale in the PISA 2006 database, thus allowing for a trend comparison between PISA 2006 and PISA 2015. Higher values on the index reflect greater levels of agreement with these statements.

Sense of belonging

The index of sense of belonging (BELONG) was constructed using students' responses to a trend question about their sense of belonging to school. Students reported, on a four-point Likert scale with the answering categories "strongly agree", "agree", "disagree", and "strongly disagree", their agreement with the following statements (ST034): 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 and out of place in my school; Other students seem to like me; I feel lonely at school. The answers to three items were reversed-coded so that higher values in the index indicate a greater sense of belonging.

Science learning in school

PISA 2015 focused on science learning in school by including several questions about the learning environment in science lessons. They asked how often specific activities happened in the school science course. The questions were used to create the following indices: teacher-directed instruction, perceived feedback, adaptive instruction, enquiry-based instruction, teacher support to students and disciplinary climate. Higher values in these indices indicate that the activities happened more frequently in science lessons.

Teacher-directed instruction

The index of teacher-directed instruction (TDTEACH) was constructed from students' reports on how often ("never or almost never"; "some lessons"; "many lessons"; "every lesson or almost every lesson") the following happened in their science lessons (ST103): The teacher explains scientific ideas; A whole class discussion takes place with the teacher; The teacher discusses our questions; The teacher demonstrates an idea.

Perceived feedback

The index of perceived feedback (PERFEED) was constructed from students' reports on how often ("never or almost never"; "some lessons"; "many lessons"; "every lesson or almost every lesson") the following happened in their science lessons (ST104): The teacher tells me how I am performing in this course; The teacher gives me feedback on my strengths in this <school science> subject; The teacher tells me in which areas I can still improve; The teacher tells me how I can improve my performance; The teacher advises me on how to reach my learning goals.

Adaptive instruction

The index of adaptive instruction (ADINST) was constructed from students' reports on how often ("never or almost never"; "some lessons"; "many lessons"; "every lesson or almost every lesson") the following happened in their science lessons (ST107): The teacher adapts the lesson to my class's needs and knowledge; The teacher provides individual help when a student has difficulties understanding a topic or task; The teacher changes the structure of the lesson on a topic that most students find difficult to understand.

Enquiry-based instruction

The index of enquiry-based instruction (IBTEACH) was constructed from students' reports on how often ("in all lessons"; "in most lessons"; "in some lessons"; "never or hardly ever") the following happened in their science lessons (ST098): Students are given opportunities to explain their ideas; Students spend time in the laboratory doing practical experiments; Students are required to argue about science questions; Students are asked to draw conclusions from an experiment they have conducted; The teacher explains how a science idea can be applied to a number of different phenomena; Students are allowed to design their own experiments; There is a class debate about investigations; The teacher clearly explains the relevance of science concepts to our lives; Students are asked to do an investigation to test ideas.

Disciplinary climate

The index of disciplinary climate (DISCLISCI) was constructed from students' reports on how often ("every lesson", "most lessons", "some lessons", "never or hardly ever") the following happened in their science lessons (ST097): The teacher shows an interest in every student's learning; The teacher gives extra help when students need it; The teacher helps students with their learning; The teacher continues teaching until students understand the material; The teacher gives students an opportunity to express their opinions. Schools were classified with having a positive disciplinary climate if the index of disciplinary climate for the school is above the national average and classified as having a negative disciplinary climate if below the national average.

Science self-efficacy

The index of science self-efficacy (SCIEEFF) was constructed based on a trend question (ST129) that was taken from PISA 2006 (ID in 2006: ST17). Students were asked, using a four-point answering scale with the categories "I could do this easily", "I could do this with a bit of effort", "I would struggle to do this on my own", and "I couldn't do this", to rate how they would perform in the following science tasks: recognise the science question that underlies a newspaper report on a health issue; explain why earthquakes occur more frequently in some areas than in others; describe the role of antibiotics in the treatment of disease; identify the science question associated with the disposal of garbage; predict how changes to an environment will affect the survival of certain species; interpret the scientific information provided on the labelling of food items; discuss how new evidence can lead you to change your understanding about the possibility of life on Mars; and identify the better of two explanations for the formation of acid rain. Responses were reverse-coded so that higher values of the index correspond to higher levels of science self-efficacy. The derived variable SCIEEFF was equated to the corresponding scale in the PISA 2006 database, thus allowing for a trend comparison between PISA 2006 and PISA 2015.

Scaling of indices related to the PISA index of economic social and cultural status

The PISA index of economic, social and cultural status (ESCS) was derived, as in previous cycles, from three variables related to family background: parents' highest level of education (PARED), parents' highest occupation status (HISEI), and home possessions (HOMEPOS), including books in the home. PARED and HISEI are simple indices, described above. HOMEPOS is a proxy measure for family wealth.

Household possessions

In PISA 2015, students reported the availability of 16 household items at home (ST011) including three country-specific household items that were seen as appropriate measures of family wealth within the country's context. In addition, students reported the amount of possessions and books at home (ST012 and ST013).

HOMEPOS is a summary index of all household and possession items (ST011, ST012 and ST013). The home possessions scale for PISA 2015 was computed differently than in the previous cycles, to align the IRT model to the one used for all cognitive and non-cognitive scales. Categories for the number of books in the home are unchanged in PISA 2015. The ST011 items (1 = "yes", 2 = "no") were reverse-coded so that a higher level indicates the presence of the indicator.

Family wealth

In PISA 2015, students reported the availability at home of a link to the Internet and a room of their own. They also reported the number of number of televisions, cars, rooms with a bath or shower, smartphones, computers (desktop computer, portable

laptop, or notebook), tablet computers, e-book readers, they have at home. In addition, countries added three specific household items that were seen as appropriate measures of family wealth within the country's context. The index of family wealth was derived from this information.

Computation of ESCS

For the purpose of computing the PISA index of economic, social and cultural status (ESCS), values for students with missing PARED, HISEI or HOMEPOS were imputed with predicted values plus a random component based on a regression on the other two variables. If there were missing data on more than one of the three variables, ESCS was not computed and a missing value was assigned for ESCS.

The PISA index of economic, social and cultural status 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. All countries and economies (both OECD and partner countries/economies) contributed equally to the principal component analysis, while in previous cycles, the principal component analysis was based on OECD countries only. However, for the purpose of reporting the ESCS scale has been transformed with zero being the score of an average OECD student and one being the standard deviation across equally weighted OECD countries.

Principal component analysis was also performed for each participating country or economy separately, to determine to what extent the components of the index operate in similar ways across countries or economy.

School-level simple indices

School type

Schools are classified as either public or private according to whether a private entity or a public agency has the ultimate power for decision making concerning its affairs (SC013). As in previous PISA surveys, the index on school type (SCHLTYPE) has three categories, based on two questions: SC013 which asks if the school is a public or a private school, and SC016 which asks about the sources of funding. This index was calculated in 2015 and in all previous cycles.

Year of reference for the trends in resources, policies and practices

Resources, policies and practices are compared between PISA 2015 and previous PISA cycles throughout the report. Whenever possible, the report compares PISA 2015 to PISA 2006 since science was the core subject in both cycles. However, PISA 2015 is compared to more recent cycles when the questions were not included in the PISA 2006 questionnaires, the wording of the questions changed (even slightly), or the number/order of the items within each question changed substantively between cycles.

Proportion of missing observations for variables used in this volume

Unless otherwise indicated, no adjustment is made for non-response to questionnaires in analyses included in this volume. The reported percentages and estimates based on indices refer to the proportion of the sample with valid responses to the corresponding questionnaire items. Tables A1.8a, A1.8b and A1.8c, available online, report the proportion of the sample covered by analyses based on student or school questionnaire variables. Where this proportion shows large variation across countries/economies or across time, caution is required when comparing results on these dimensions. Table A1.8d reports the differences in student characteristics between students with available data and students with missing data.

Derivation of the index of exposure to bullying

The development of comparable measures of student and school characteristics from the student and school questionnaires is a major goal of PISA. Cross-country validity of the measured items requires more than a thorough process of translation into different languages. It also makes assumptions about having measured similar characteristics in different national and cultural contexts. Many questionnaire items in PISA are designed to be combined in some way in order to measure latent constructs that cannot be observed directly (e.g. a student's achievement motivation). Transformations or scaling procedures are applied to these items in order to construct meaningful indices (OECD, forthcoming).

PISA 2015 includes eight items on students' exposure to bullying or bullying victimisation. A scale for exposure to bullying is not included in the international database, but was derived for this report using confirmatory factor analysis (CFA). This annex describes how the scale was constructed and reports the results of tests of the measurement reliability and cross-country invariance of the scale. These tests are important because international comparisons and analysis based on the scale are possible as long as the latent construct ("exposure to bullying" in this case) is the same and measured in the same way across different countries and economies. The scaling analysis used the software Mplus, Version 7.1 (Muthén and Muthén, 1998-2012).

Exploratory analysis of the data showed that the first two of the eight items on bullying did not load well onto a unidimensional construct and were also not strongly correlated with the other six items. The averages of these two items also vary across countries much more than the other six items, potentially indicating measurement issues (e.g. students in some countries might have interpreted the questions differently from students in other countries). In order to produce a scale of bullying with a sufficiently good model fit in all countries and comparability across countries, the scaling was limited to the six other items. Students reported how frequently they were exposed to the types of bullying described by the six items, according to

a four-point scale: 1) "Never or almost never"; 2) "A few times a year"; 3) "A few times a month"; 4) "Once a week or more". In alignment with how previous literature has defined "frequent bullying" [Salmivalli et al., 2011], categories 3) and 4) were aggregated into a single category. Such aggregation only marginally affected the overall fit of the scale but improved the international invariance of the scale. Students might find it relatively difficult to distinguish between "a few times a month" and "once a week or more", so that the variation between the two categories might reflect different interpretations of the question or different response styles across countries, rather than real differences in exposure to bullying. Figure A1.1 summarises how the original data in PISA 2015 were selected and recoded for scaling purposes.

Figure A1.1 ■ **Questionnaire items used for the scale of exposure to bullying**

	"Never or almost never"	"A few times a year"	"A few times a month" or "Once a week or more" (the two categories are merged)
~~Q01: I got called names by other students.~~ (Not used for the scale)	~~☐1~~	~~☐2~~	~~☐3~~
~~Q02: I got picked on by other students.~~ (Not used for the scale)	~~☐1~~	~~☐2~~	~~☐3~~
Q03: Other students left me out of things on purpose.	☐1	☐2	☐3
Q04: Other students made fun of me.	☐1	☐2	☐3
Q05: I was threatened by other students.	☐1	☐2	☐3
Q06: Other students took away or destroyed things that belonged to me.	☐1	☐2	☐3
Q07: I got hit or pushed around by other students.	☐1	☐2	☐3
Q08: Other students spread nasty rumours about me.	☐1	☐2	☐3

More frequent bullying

The data on bullying are not continuous but take one of the three frequency categories, and thus require a model that explicitly accounts for this categorical distribution (Muthén, 1997, 1993). The model assumes that an observed variable, *x* (one of the six types of bullying), comes from a latent response variable, *x** (the student's actual exposure to that type of bullying). The observed categories of *x* for each student *i* correspond to particular thresholds along the continuum of the latent variable *x**:

x_i = "Never or almost never"(category 1) if $x_i^* \leq \tau_{i,1}$;

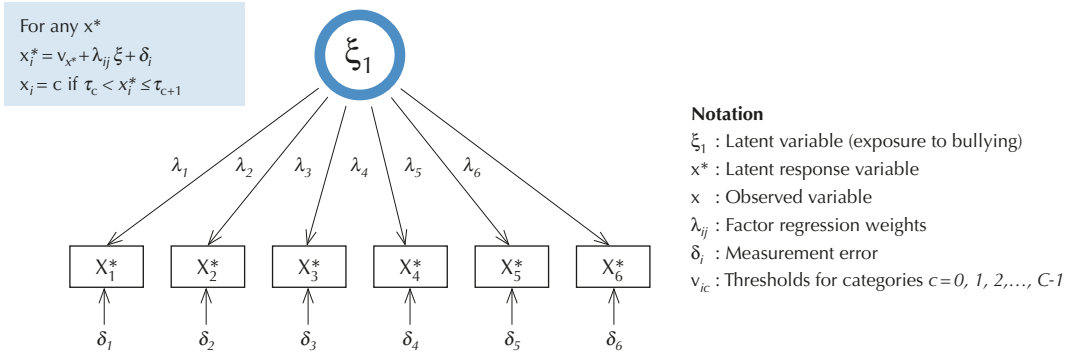
x_i = "A few times a year" (category 2) if $\tau_{i,1} < x_i^* \leq \tau_{i,2}$;

And

x_i = "A few times a month" or "once a week or more" (category 3); if $x_i^* > \tau_{i,2}$

The thresholds are parameters to be estimated in the model. Figure A1.2 provides a graphical representation of the model used to scale the six items on bullying. The model uses a theta parameterisation and fixes for identification the first factor loading to 1, the latent variable means to 0 and the residual variance to 1 across all groups.

Figure A1.2 ■ **Representation of the categorical model for the scale of exposure to bullying**



A first method to check the reliability of the scale is to estimate the correlation between the different items included in the scale across all countries. Cronbach's alpha measures the average covariance between item pairs, and can be used to check the internal consistency of a scaled index within the countries and to compare it between the countries (OECD, forthcoming). Table A1.4 shows that, on average (assigning equal weight to all countries with available data) the scale of exposure to bullying has a Cronbach alpha of 0.83. The Cronbach alpha ranges between 0.71 (lowest) for Korea to 0.9 (highest) for Qatar, suggesting that the correlation between the six items included in the scale is acceptable in most countries.

Measurement invariance of the scale is usually established through a set of hierarchical tests, ranging from least strict to most strict. Chi-square tests, chi-square difference tests, fit indices, and changes in fit indices across specifications are typical measures of measurement invariance. Three levels of invariance are analysed in this annex: 1) configural (or baseline) invariance; 2) metric (or equal slopes) invariance; 3) scalar (or equal slopes and thresholds) invariance. Configural invariance is verified if, for two or more populations, the same construct is measured with the same indicators in the same way. Metric invariance requires that, in addition to configural invariance, all factor loadings are statistically equivalent. For scalar invariance, in addition to metric invariance, all thresholds should be statistically equivalent.

When the slope and thresholds for all items in the measurement model are not significantly different across groups, full scalar equivalence is achieved. However, Byrne et al. (1989) have argued that full scalar equivalence is not a necessary condition for comparisons to be valid. If at least two items per latent variable (namely, the item that is fixed at unity to identify the model and one other item) are equivalent, comparisons can be validly made across countries (Steenkamp and Baumgartner, 1998). Thus, partial equivalence does not require the invariance of all loadings and intercepts in all countries. The final model used for the bullying scale was based on a partial-invariance specification in which at least three items are fixed across all countries, and up to three items are allowed to vary across 11 countries and economies (see Table A1.5 for details on which constraints were relaxed in which countries). The selection of the country-items pairs that were freely estimated was determined empirically, on the basis on the deterioration of fit associated with constraining these items to baseline values.

Table A1.5 reports the contribution of the different countries/economies to the Chi-square fit statistic under three different model specifications (configural, scalar and scalar with partial invariance). A high value of the Chi-square test statistic indicates a worse fit of the model. The Chi-square is sensitive to sample size (Bentler and Bonett, 1980).

Table A1.3 shows the change in model fit associated with assuming metric and scalar invariance, under the full and partial invariance specifications. The model fit is measured by the comparative fit index (CFI) and by the Root Mean Square Error of Approximation (RMSEA). A value of CFI equal to 1 indicates perfect fit; a value around 0.9 is generally considered acceptable. A value of the RMSEA equal to 0.00 indicates perfect fit; values between 0.05 and 0.08 are considered acceptable. As can be seen from the table, allowing up to three items to be estimated freely in a limited number of countries significantly reduces the deterioration in the model fit associated with assuming equal slopes in all countries. When allowing factor loadings to vary for up to 3 items in 11 countries and economies, the change in the model fit is within defensible criteria for measurement invariance in categorical models (Rutkowski and Svetina, 2017; Rutkowski and Svetina, 2013). These findings support, to some extent, the international comparisons described in Chapter 8. However, given that only partial and not full invariance could be verified, some caution needs to be exercised in interpreting cross-country analysis based on this scale.

Table A1.3 ■ **Change in fit indexes with restrictions for full and partial invariance**

	Configural	Metric	Scalar	Change in model fit (Metric – Configural)	Change in model fit (Scalar-Configural)
	Full invariance				
CFI[1]	0.989	0.978	0.979	-0.011	-0.010
RMSEA[2]	0.069	0.066	0.076	-0.003	0.007
	Partial invariance				
CFI	0.989	0.984	0.982	-0.005	-0.007
RMSEA	0.069	0.068	0.061	-0.001	-0.008

1. Comparative Fit Index.
2. Root Mean Square Error Of Approximation.

Tables A1.6a, A1.6b and A1.6c report the factor loadings and thresholds for the baseline model (configural) and for the specification with partially fixed slopes and thresholds (scalar) that accounts for PISA's complex sampling scheme. Table A1.7 shows the rate of victimisation by item of students in the top 10% of the international index of exposure to bullying.

Table available online

Table A1.4 Cronbach Alpha reliability coefficients for the scale of exposure to bullying
(http://dx.doi.org/10.1787/888933473532)

Table A1.5 Chi-Square tests of model fit
(http://dx.doi.org/10.1787/888933473544)

Table A1.6a Factor loadings for the configural (baseline) model
(http://dx.doi.org/10.1787/888933473558)

Table A1.6b Factor loadings for the scalar model with partial invariance and replicate weights
(http://dx.doi.org/10.1787/888933473565)

Table A1.6c Estimated thresholds for the configural (baseline) model
(http://dx.doi.org/10.1787/888933473578)

Table A1.6d Estimated thresholds for the scalar model with partial invariance and replicate weights
(http://dx.doi.org/10.1787/888933473585)

Table A1.7 Rate of victimisation of "frequently bullied students"
(http://dx.doi.org/10.1787/888933473597)

Table A1.8a Weighted share of responding students covered by analyses based on student and educational career questionnaire
(http://dx.doi.org/10.1787/888933473606)

Table A1.8b Weighted share of responding students covered by analyses based on school questionnaire
(http://dx.doi.org/10.1787/888933473611)

Table A1.8c Weighted share of responding students covered by analyses based on parent questionnaire
(http://dx.doi.org/10.1787/888933473622)

Table A1.8d Differences between students with complete and students with missing observations on the parental questionnaire
(http://dx.doi.org/10.1787/888933473637)

References

Bentler, P. M. and **D.G. Bonett** (1980), "Significance tests and goodness of fit in the analysis of covariance structures", *Psychological Bulletin*, Vol. 88/3, pp. 588-606, http://dx.doi.org/10.1037/0033-2909.88.3.588.

Byrne, B. M., R.J. Shavelson and **B. Muthén** (1989), "Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance", *Psychological Bulletin*, Vol. 105/3, pp. 456-466, http://dx.doi.org/10.1037/0033-2909.105.3.456.

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", Conference paper presented at the *Annual Conference of International Social Survey Programme*, Lisbon, Portugal.

Ganzeboom, H. B.G. and **D.J. Treiman** (2003), "Three internationally standardised measures for comparative research on occupational status", in J.H.P. Hoffmeyer-Zlotnik and C. Wolf (eds.), *Advances in Cross-National Comparison: A European Working Book for Demographic and Socio-Economic Variables*, Kluwer Academic Press, New York, NY, pp. 159-193.

Kjærnsli, M. and **S. Lie** (2011), "Students' preference for science careers: International comparisons based on PISA 2006", *International Journal of Science Education*, Vol. 33/1, pp. 121-44, http://dx.doi.org/10.1080/09500693.2010.518642.

OECD (forthcoming), *PISA 2015 Technical Report,* PISA, OECD Publishing, Paris.

OECD (2016), *PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematic and Financial Literacy,* PISA, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264255425-en.

OECD (2007), *PISA 2006: Science Competencies for Tomorrow's World*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264040014-en.

OECD (1999), *Classifying Educational Programmes: Manual for ISCED-97 Implementation in OECD Countries,* OECD Publishing, Paris.

O*NET OnLine (n.d), "All STEM disciplines", web page, www.onetonline.org/find/stem?t=0, (accessed 4 October 2016).

Rutkowski, L. and **D. Svetina** (2017), "Measurement invariance in international surveys: Categorical indicators and fit measure performance", *Applied Measurement in Education*, Vol. 30/1, pp. 39-51,

Rutkowski, L. and **D. Svetina** (2013), "Assessing the hypothesis of measurement invariance in the context of large-scale international surveys", *Educational and Psychological Measurement*, Vol. 74/1, pp. 31-57, http://dx.doi.org/110.1177/0013164413498257.

Sikora, J. and **A. Pokropek** (2012), "Gender segregation of adolescent science career plans in 50 countries", *Science Education*, Vol. 96/2, pp. 234-264, http://dx.doi.org/10.1002/sce.20479.

Salmivalli C., A. Kärnä and **E. Poskiparta** (2011), "Counteracting bullying in Finland: The KiVa Program and its effects on different forms of being bullied", *International Journal of Behavioral Development*, Vol. 35, pp. 405-411, http://doi.org/10.1177/0165025411407457.

Steenkamp, J. and **H. Baumgartner** (1998), "Assessing measurement invariance in cross-national consumer research", *Journal of Consumer Research*, Vol. 25/1, pp. 78-107, http://doi.org/10.1086/209528.

Warm, T.A. (1989), "Weighted likelihood estimation of ability in item response theory", *Psychometrika*, Vol. 54/3, pp. 427-450, http://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 2015 provides an assessment of the cumulative outcomes 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.

Differences between countries in the nature and extent of pre-primary education and care, the age at entry into formal schooling and the institutional structure of education systems do not allow for a definition of internationally comparable grade levels. Consequently, international comparisons of performance in education 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 level 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 18 days (0.20 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 school. In PISA, these knowledge and skills are referred to as the outcomes 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 later on as/if students' educational experiences converge over time.

If a country's scores in science, reading or mathematics 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 in school, home and beyond, have resulted in higher outcomes in the literacy domains that PISA measures.

The PISA target population does 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 requested grade-based results for the purpose of national analyses, PISA 2015 provided a sampling option to supplement age-based sampling with grade-based sampling.

Population coverage

All countries and economies attempted to maximise the coverage of 15-year-olds enrolled in education in their national samples, including students enrolled in special-education institutions. As a result, PISA 2015 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 12 countries – the United Kingdom (8.22%), Luxembourg (8.16%), Canada (7.49%), Norway (6.75%), New Zealand (6.54%), Sweden (5.71%), Estonia (5.52%), Australia (5.31%),

Montenegro (5.17%), Lithuania (5.12%), Latvia (5.07%), and Denmark (5.04%) – achieved this standard, and in 29 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), Denmark, Latvia, New Zealand and Sweden 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: schools that were geographically inaccessible or where the administration of the PISA assessment was not considered feasible; and 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 the former group and 2% maximum for the latter group). The magnitude, nature and justification of school-level exclusions are documented in the *PISA 2015 Technical Report* (OECD, forthcoming).
- At the student level: students with an intellectual disability; students with a functional disability; students with limited assessment language proficiency; other (a category defined by the national centres and approved by the international centre); and 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 disciplinary 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 2015. Further information on the target population and the implementation of PISA sampling standards can be found in the *PISA 2015 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 means the year 2014 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 0.21% of its population for a particular type of student educated while working; Canada excluded 1.22% 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; and the United Arab Emirates excluded 0.04% of its students who had no information available. The adjudicated region of Massachusetts in the United States excluded 13.11% of its students, and North Carolina excluded 5.64% of its students. For these two regions, the desired target populations cover 15-year-old students in grade 7 or above in public schools only. The students excluded from the desired population are private school students.
- **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 2015. 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 PISA's 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: 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); 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); students with limited proficiency in the assessment language (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 assessment may be excluded); other (a category defined by the national centres and approved by the international centre); and students taught in a language of instruction for the main domain for which no materials were available.

[Part 1/1]

Table A2.1 PISA target populations and samples

		Population and sample information												Coverage indices		
		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	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
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
OECD	Australia	282 888	282 547	282 547	6 940	275 607	2.46	14 530	256 329	681	7 736	2.93	5.31	0.947	0.947	0.906
	Austria	88 013	82 683	82 683	790	81 893	0.96	7 007	73 379	84	866	1.17	2.11	0.979	0.979	0.834
	Belgium	123 630	121 954	121 694	1 597	120 097	1.31	9 651	114 902	39	410	0.36	1.66	0.983	0.981	0.929
	Canada	396 966	381 660	376 994	1 590	375 404	0.42	20 058	331 546	1 830	25 340	7.10	7.49	0.925	0.914	0.835
	Chile	255 440	245 947	245 852	2 641	243 211	1.07	7 053	203 782	37	1 393	0.68	1.75	0.983	0.982	0.798
	Czech Republic	90 391	90 076	90 076	1 814	88 262	2.01	6 894	84 519	25	368	0.43	2.44	0.976	0.976	0.935
	Denmark	68 174	67 466	67 466	605	66 861	0.90	7 161	60 655	514	2 644	4.18	5.04	0.950	0.950	0.890
	Estonia	11 676	11 491	11 491	416	11 075	3.62	5 587	10 834	116	218	1.97	5.52	0.945	0.945	0.928
	Finland	58 526	58 955	58 955	472	58 483	0.80	5 882	56 934	124	1 157	1.99	2.78	0.972	0.972	0.973
	France	807 867	778 679	778 679	28 742	749 937	3.69	6 108	734 944	35	3 620	0.49	4.16	0.958	0.958	0.910
	Germany	774 149	774 149	774 149	11 150	762 999	1.44	6 522	743 969	54	5 342	0.71	2.14	0.979	0.979	0.961
	Greece	105 530	105 253	105 253	953	104 300	0.91	5 532	96 157	58	965	0.99	1.89	0.981	0.981	0.911
	Hungary	94 515	90 065	90 065	1 945	88 120	2.16	5 658	84 644	55	1 009	1.18	3.31	0.967	0.967	0.896
	Iceland	4 250	4 195	4 195	17	4 178	0.41	3 374	3 966	131	132	3.23	3.62	0.964	0.964	0.933
	Ireland	61 234	59 811	59 811	72	59 739	0.12	5 741	59 082	197	1 825	3.00	3.11	0.969	0.969	0.965
	Israel	124 852	118 997	118 997	2 310	116 687	1.94	6 598	117 031	115	1 803	1.52	3.43	0.966	0.966	0.937
	Italy	616 761	567 268	567 268	11 190	556 078	1.97	11 583	495 093	246	9 395	1.86	3.80	0.962	0.962	0.803
	Japan	1 201 615	1 175 907	1 175 907	27 323	1 148 584	2.32	6 647	1 138 349	2	318	0.03	2.35	0.976	0.976	0.947
	Korea	620 687	619 950	619 950	3 555	616 395	0.57	5 581	569 106	20	1 806	0.32	0.89	0.991	0.991	0.917
	Latvia	17 255	16 955	16 955	677	16 278	3.99	4 869	15 320	70	174	1.12	5.07	0.949	0.949	0.888
	Luxembourg	6 327	6 053	6 053	162	5 891	2.68	5 299	5 540	331	331	5.64	8.16	0.918	0.918	0.876
	Mexico	2 257 399	1 401 247	1 401 247	5 905	1 395 342	0.42	7 568	1 392 995	30	6 810	0.49	0.91	0.991	0.991	0.617
	Netherlands	201 670	200 976	200 976	6 866	194 110	3.42	5 385	191 817	14	502	0.26	3.67	0.963	0.963	0.951
	New Zealand	60 162	57 448	57 448	681	56 767	1.19	4 520	54 274	333	3 112	5.42	6.54	0.935	0.935	0.902
	Norway	63 642	63 491	63 491	854	62 637	1.35	5 456	58 083	345	3 366	5.48	6.75	0.933	0.933	0.913
	Poland	380 366	361 600	361 600	6 122	355 478	1.69	4 478	345 709	34	2 418	0.69	2.38	0.976	0.976	0.909
	Portugal	110 939	101 107	101 107	424	100 683	0.42	7 325	97 214	105	860	0.88	1.29	0.987	0.987	0.876
	Slovak Republic	55 674	55 203	55 203	1 376	53 827	2.49	6 350	49 654	114	912	1.80	4.25	0.957	0.957	0.892
	Slovenia	18 078	17 689	17 689	290	17 399	1.64	6 406	16 773	114	247	1.45	3.07	0.969	0.969	0.928
	Spain	440 084	414 276	414 276	2 175	412 101	0.53	6 736	399 935	200	10 893	2.65	3.16	0.968	0.968	0.909
	Sweden	97 749	97 210	97 210	1 214	95 996	1.25	5 458	91 491	275	4 324	4.51	5.71	0.943	0.943	0.936
	Switzerland	85 495	83 655	83 655	2 320	81 335	2.77	5 860	82 223	107	1 357	1.62	4.35	0.956	0.956	0.962
	Turkey	1 324 089	1 100 074	1 100 074	5 746	1 094 328	0.52	5 895	925 366	31	5 359	0.58	1.10	0.989	0.989	0.699
	United Kingdom	747 593	746 328	746 328	23 412	722 916	3.14	14 157	627 703	870	34 747	5.25	8.22	0.918	0.918	0.840
	United States	4 220 325	3 992 053	3 992 053	12 001	3 980 052	0.30	5 712	3 524 497	193	109 580	3.02	3.31	0.967	0.967	0.835
Partners	Albania	48 610	45 163	45 163	10	45 153	0.02	5 215	40 896	0	0	0.00	0.02	1.000	1.000	0.841
	Algeria	389 315	354 936	354 936	0	354 936	0.00	5 519	306 647	0	0	0.00	0.00	1.000	1.000	0.788
	Argentina	718 635	578 308	578 308	2 617	575 691	0.45	6 349	394 917	21	1 367	0.34	0.80	0.992	0.992	0.550
	Brazil	3 430 255	2 853 388	2 853 388	64 392	2 788 996	2.26	23 141	2 425 961	119	13 543	0.56	2.80	0.972	0.972	0.707
	B-S-J-G (China)	2 084 958	1 507 518	1 507 518	58 639	1 448 879	3.89	9 841	1 331 794	33	3 609	0.27	4.15	0.959	0.959	0.639
	Bulgaria	66 601	59 397	59 397	1 124	58 273	1.89	5 928	53 685	49	433	0.80	2.68	0.973	0.973	0.806
	Colombia	760 919	674 079	674 079	37	674 042	0.01	11 795	567 848	9	507	0.09	0.09	0.999	0.999	0.746
	Costa Rica	81 773	66 524	66 524	0	66 524	0.00	6 866	51 897	13	98	0.19	0.19	0.998	0.998	0.635
	Croatia	45 031	35 920	35 920	805	35 115	2.24	5 809	40 899	86	589	1.42	3.63	0.964	0.964	0.908
	Cyprus*	9 255	9 255	9 253	109	9 144	1.18	5 571	8 785	228	292	3.22	4.36	0.956	0.956	0.949
	Dominican Republic	193 153	139 555	139 555	2 382	137 173	1.71	4 740	132 300	4	106	0.08	1.79	0.982	0.982	0.685
	FYROM	16 719	16 717	16 717	259	16 458	1.55	5 324	15 847	8	19	0.12	1.67	0.983	0.983	0.948
	Georgia	48 695	43 197	43 197	1 675	41 522	3.88	5 316	38 334	35	230	0.60	4.45	0.955	0.955	0.787
	Hong Kong (China)	65 100	61 630	61 630	708	60 922	1.15	5 359	57 662	36	374	0.65	1.79	0.982	0.982	0.886
	Indonesia	4 534 216	3 182 816	3 182 816	4 046	3 178 770	0.13	6 513	3 092 773	0	0	0.00	0.13	0.999	0.999	0.682
	Jordan	126 399	121 729	121 729	71	121 658	0.06	7 267	108 669	70	1 006	0.92	0.97	0.990	0.990	0.860
	Kazakhstan	211 407	209 555	209 555	7 475	202 080	3.57	7 841	192 909	0	0	0.00	3.57	0.964	0.964	0.912
	Kosovo	31 546	28 229	28 229	1 156	27 073	4.10	4 826	22 333	50	174	0.77	4.84	0.952	0.952	0.708
	Lebanon	64 044	62 281	62 281	1 300	60 981	2.09	4 546	42 331	0	0	0.00	2.09	0.979	0.979	0.661
	Lithuania	33 163	32 097	32 097	573	31 524	1.79	6 525	29 915	227	1 050	3.39	5.12	0.949	0.949	0.902
	Macao (China)	5 100	4 417	4 417	3	4 414	0.07	4 476	4 507	0	0	0.00	0.07	0.999	0.999	0.884
	Malaysia	540 000	448 838	448 838	2 418	446 420	0.54	8 861	412 524	41	2 344	0.56	1.10	0.989	0.989	0.764
	Malta	4 397	4 406	4 406	63	4 343	1.43	3 634	4 296	41	41	0.95	2.36	0.976	0.976	0.977
	Moldova	31 576	30 601	30 601	182	30 419	0.59	5 325	29 341	21	118	0.40	0.99	0.990	0.990	0.929
	Montenegro	7 524	7 506	7 506	40	7 466	0.53	5 665	6 777	300	332	4.66	5.17	0.948	0.948	0.901
	Peru	580 371	478 229	478 229	6 355	471 874	1.33	6 971	431 738	13	745	0.17	1.50	0.985	0.985	0.744
	Qatar	13 871	13 850	13 850	380	13 470	2.74	12 083	12 951	193	193	1.47	4.17	0.958	0.958	0.934
	Romania	176 334	176 334	176 334	1 823	174 511	1.03	4 876	164 216	3	120	0.07	1.11	0.989	0.989	0.931
	Russia	1 176 473	1 172 943	1 172 943	24 217	1 148 726	2.06	6 036	1 120 932	13	2 469	0.22	2.28	0.977	0.977	0.953
	Singapore	48 218	47 050	47 050	445	46 605	0.95	6 115	46 224	25	179	0.39	1.33	0.987	0.987	0.959
	Chinese Taipei	295 056	287 783	287 783	1 179	286 604	0.41	7 708	251 424	22	647	0.26	0.67	0.993	0.993	0.852
	Thailand	895 513	756 917	756 917	9 646	747 271	1.27	8 249	634 795	22	2 107	0.33	1.60	0.984	0.984	0.709
	Trinidad and Tobago	17 371	17 371	17 371	0	17 371	0.00	4 692	13 197	0	0	0.00	0.00	1.000	1.000	0.760
	Tunisia	122 186	122 186	122 186	679	121 507	0.56	5 375	113 599	3	61	0.05	0.61	0.994	0.994	0.930
	United Arab Emirates	51 687	51 518	51 499	994	50 505	1.93	14 167	46 950	63	152	0.32	2.25	0.978	0.977	0.908
	Uruguay	53 533	43 865	43 865	4	43 861	0.01	6 062	38 287	6	32	0.08	0.09	0.999	0.999	0.715
	Viet Nam	1 803 552	1 032 599	1 032 599	6 557	1 026 042	0.63	5 826	874 859	0	0	0.00	0.63	0.994	0.994	0.485

Notes: For a full explanation of the details in this table please refer to the *PISA 2015 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.
For Mexico, in 2015, the Total population of 15-year-olds enrolled in grade 7 or above is an estimate of the target population size of the sample frame from which the 15-year-olds students were selected for the PISA test. At the time Mexico provided the information to PISA, the official figure for this population was 1 573 952.
* See note at the beginning of this Annex.
StatLink http://dx.doi.org/10.1787/888933433129

[Part 1/2]

Table A2.2 Exclusions

		Student exclusions (unweighted)					
		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)	School-level exclusion rate (%)
		(1)	(2)	(3)	(4)	(5)	(6)
OECD	Australia	85	528	68	0	0	681
	Austria	8	15	61	0	0	84
	Belgium	4	18	17	0	0	39
	Canada	156	1 308	366	0	0	1 830
	Chile	6	30	1	0	0	37
	Czech Republic	2	9	14	0	0	25
	Denmark	18	269	156	70	1	514
	Estonia	17	93	6	0	0	116
	Finland	2	90	17	8	7	124
	France	5	21	9	0	0	35
	Germany	4	25	25	0	0	54
	Greece	3	44	11	0	0	58
	Hungary	3	13	9	30	0	55
	Iceland	9	66	47	9	0	131
	Ireland	25	57	55	60	0	197
	Israel	22	68	25	0	0	115
	Italy	78	147	21	0	0	246
	Japan	0	2	0	0	0	2
	Korea	3	17	0	0	0	20
	Latvia	7	47	16	0	0	70
	Luxembourg	4	254	73	0	0	331
	Mexico	4	23	3	0	0	30
	Netherlands	1	13	0	0	0	14
	New Zealand	23	140	167	0	3	333
	Norway	11	253	81	0	0	345
	Poland	11	20	0	3	0	34
	Portugal	4	99	2	0	0	105
	Slovak Republic	7	71	2	34	0	114
	Slovenia	33	36	45	0	0	114
	Spain	9	144	47	0	0	200
	Sweden	154	0	121	0	0	275
	Switzerland	8	42	57	0	0	107
	Turkey	1	23	7	0	0	31
	United Kingdom	77	690	102	0	1	870
	United States	16	120	44	13	0	193
Partners	Albania	0	0	0	0	0	0
	Algeria	0	0	0	0	0	0
	Argentina	10	10	1	0	0	21
	Brazil	20	99	0	0	0	119
	B-S-J-G (China)	6	25	2	0	0	33
	Bulgaria	39	6	4	0	0	49
	Colombia	3	4	2	0	0	9
	Costa Rica	3	1	0	9	0	13
	Croatia	2	75	9	0	0	86
	Cyprus*	12	164	52	0	0	228
	Dominican Republic	1	3	0	0	0	4
	FYROM	7	1	0	0	0	8
	Georgia	3	25	7	0	0	35
	Hong Kong (China)	0	35	1	0	0	36
	Indonesia	0	0	0	0	0	0
	Jordan	43	17	10	0	0	70
	Kazakhstan	0	0	0	0	0	0
	Kosovo	9	13	27	0	0	50
	Lebanon	0	0	0	0	0	0
	Lithuania	12	213	2	0	0	227
	Macao (China)	0	0	0	0	0	0
	Malaysia	10	22	9	0	0	41
	Malta	8	27	6	0	0	41
	Moldova	12	8	1	0	0	21
	Montenegro	14	23	5	0	258	300
	Peru	4	9	0	0	0	13
	Qatar	76	110	7	0	0	193
	Romania	1	1	1	0	0	3
	Russia	3	10	0	0	0	13
	Singapore	3	15	7	0	0	25
	Chinese Taipei	3	19	0	0	0	22
	Thailand	1	19	2	0	0	22
	Trinidad and Tobago	0	0	0	0	0	0
	Tunisia	0	0	3	0	0	3
	United Arab Emirates	16	24	23	0	0	63
	Uruguay	2	4	0	0	0	6
	Viet Nam	0	0	0	0	0	0

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 2015 Technical Report* (OECD, forthcoming).

* See note at the beginning of this Annex.
StatLink http://dx.doi.org/10.1787/888933433129

[Part 2/2]

Table A2.2 Exclusions

		Student exclusion (weighted)					
		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)	Weighted number of excluded students because of no materials available in the language of instruction (Code 5)	Total weighted number of excluded students
		(7)	(8)	(9)	(10)	(11)	(12)
OECD	Australia	932	6 011	793	0	0	7 736
	Austria	74	117	675	0	0	866
	Belgium	33	192	185	0	0	410
	Canada	1 901	18 018	5 421	0	0	25 340
	Chile	194	1 190	9	0	0	1 393
	Czech Republic	40	140	188	0	0	368
	Denmark	122	1 539	551	421	11	2 644
	Estonia	29	176	13	0	0	218
	Finland	18	858	156	67	58	1 157
	France	562	2 144	914	0	0	3 620
	Germany	423	2 562	2 357	0	0	5 342
	Greece	43	729	193	0	0	965
	Hungary	57	284	114	554	0	1 009
	Iceland	9	67	47	9	0	132
	Ireland	213	526	516	570	0	1 825
	Israel	349	1 070	384	0	0	1 803
	Italy	3 316	5 199	880	0	0	9 395
	Japan	0	318	0	0	0	318
	Korea	291	1 515	0	0	0	1 806
	Latvia	21	115	38	0	0	174
	Luxembourg	4	254	73	0	0	331
	Mexico	842	4 802	1 165	0	0	6 810
	Netherlands	33	469	0	0	0	502
	New Zealand	233	1 287	1 568	0	24	3 112
	Norway	105	2 471	790	0	0	3 366
	Poland	876	1 339	0	203	0	2 418
	Portugal	29	818	13	0	0	860
	Slovak Republic	44	567	12	288	0	912
	Slovenia	84	71	92	0	0	247
	Spain	511	7 662	2 720	0	0	10 893
	Sweden	2 380	0	1 944	0	0	4 324
	Switzerland	91	540	726	0	0	1 357
	Turkey	43	4 094	1 222	0	0	5 359
	United Kingdom	2 724	27 808	4 001	0	214	34 747
	United States	7 873	67 816	26 525	7 366	0	109 580
Partners	Albania	0	0	0	0	0	0
	Algeria	0	0	0	0	0	0
	Argentina	579	770	18	0	0	1 367
	Brazil	1 743	11 800	0	0	0	13 543
	B-S-J-G (China)	438	2 970	201	0	0	3 609
	Bulgaria	347	51	35	0	0	433
	Colombia	181	309	17	0	0	507
	Costa Rica	22	5	0	71	0	98
	Croatia	13	501	75	0	0	589
	Cyprus*	16	212	65	0	0	292
	Dominican Republic	24	82	0	0	0	106
	FYROM	15	4	0	0	0	19
	Georgia	19	170	41	0	0	230
	Hong Kong (China)	0	363	11	0	0	374
	Indonesia	0	0	0	0	0	0
	Jordan	656	227	122	0	0	1 006
	Kazakhstan	0	0	0	0	0	0
	Kosovo	28	37	104	0	0	174
	Lebanon	0	0	0	0	0	0
	Lithuania	40	1 000	10	0	0	1 050
	Macao (China)	0	0	0	0	0	0
	Malaysia	663	1 100	580	0	0	2 344
	Malta	8	27	6	0	0	41
	Moldova	66	51	1	0	0	118
	Montenegro	27	38	6	0	261	332
	Peru	224	520	0	0	0	745
	Qatar	76	110	7	0	0	193
	Romania	31	63	26	0	0	120
	Russia	425	2 044	0	0	0	2 469
	Singapore	22	115	43	0	0	179
	Chinese Taipei	78	568	0	0	0	647
	Thailand	114	1 830	163	0	0	2 107
	Trinidad and Tobago	0	0	0	0	0	0
	Tunisia	0	0	61	0	0	61
	United Arab Emirates	30	75	47	0	0	152
	Uruguay	10	22	0	0	0	32
	Viet Nam	0	0	0	0	0	0

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 2015 Technical Report* (OECD, forthcoming).

* See note at the beginning of this Annex.

StatLink http://dx.doi.org/10.1787/888933433129

- **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.
- **Column 13** presents an index of the extent to which the national desired target population is covered by the PISA sample. Australia, Canada, Denmark, Estonia, Latvia, Lithuania, Luxembourg, Montenegro, New Zealand, Norway, Sweden and the United Kingdom 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 2015. 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) (times 100).
- **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 on 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 used that assumes a bivariate normal distribution for performance and the propensity to participate. For details, see the *PISA 2015 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 2015 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. At least 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 2015.

In the case of Iceland, Luxembourg, Macao (China), Malta 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, 42 students were then selected with equal probability (all 15-year-old students were selected if fewer than 42 were enrolled). The number of students to be sampled per school could deviate from 42, 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 2015 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.

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.
- **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.
- **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.
- **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 of 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, subunits 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 and Croatia, 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. In Luxembourg, a school on the border with Germany was split according to the country in which the students resided. In addition, the International schools in Luxembourg were split into the students who were instructed in any of the three official languages, and those in the part of the schools that was excluded because no materials were available in the languages of instruction. The United Arab Emirates had schools split by curricula, and sometimes by gender, with other schools remaining whole. Because of reorganisation, some of Sweden's schools were split into parts, with each part having one principal. In Portugal, schools were reorganised into clusters, with teachers and the principal shared by all units in the school cluster.

[Part 1/1]

Table A2.3 Response rates

		Initial sample – before school replacement					Final sample – after school replacement					Final sample – students within schools 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 and non-responding schools (unweighted)	Total in national desired target population after all school exclusions and before within-school exclusions	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)	Number of responding schools (unweighted)	Number of responding and 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)
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
OECD	Australia	94	260 657	276 072	720	788	95	262 130	276 072	723	788	84	204 763	243 789	14 089	17 477
	Austria	100	81 690	81 730	269	273	100	81 690	81 730	269	273	87	63 660	73 521	7 007	9 868
	Belgium	83	98 786	118 915	244	301	95	113 435	118 936	286	301	91	99 760	110 075	9 635	10 602
	Canada	74	283 853	381 133	703	1 008	79	299 512	381 189	726	1 008	81	210 476	260 487	19 604	24 129
	Chile	92	215 139	232 756	207	232	99	230 749	232 757	226	232	93	189 206	202 774	7 039	7 515
	Czech Republic	98	86 354	87 999	339	344	98	86 354	87 999	339	344	89	73 386	82 672	6 835	7 693
	Denmark	90	57 803	63 897	327	371	92	58 837	63 931	331	371	89	49 732	55 830	7 149	8 184
	Estonia	100	11 142	11 154	206	207	100	11 142	11 154	206	207	93	10 088	10 822	5 587	5 994
	Finland	100	58 653	58 782	167	168	100	58 800	58 800	168	168	93	53 198	56 934	5 882	6 294
	France	91	679 984	749 284	232	255	94	706 838	749 284	241	255	88	611 563	693 336	5 980	6 783
	Germany	96	764 423	794 206	245	256	99	785 813	794 206	253	256	93	685 972	735 487	6 476	6 944
	Greece	92	95 030	103 031	190	212	98	101 653	103 218	209	212	94	89 588	94 986	5 511	5 838
	Hungary	93	83 897	89 808	231	251	99	88 751	89 825	244	251	92	77 212	83 657	5 643	6 101
	Iceland	99	4 114	4 163	122	129	99	4 114	4 163	122	129	86	3 365	3 908	3 365	3 908
	Ireland	99	61 023	61 461	167	169	99	61 023	61 461	167	169	89	51 947	58 630	5 741	6 478
	Israel	91	105 192	115 717	169	190	93	107 570	115 717	173	190	90	98 572	108 940	6 598	7 294
	Italy	74	383 933	516 113	414	532	88	451 098	515 515	464	532	88	377 011	430 041	11 477	12 841
	Japan	94	1 087 414	1 151 305	189	200	99	1 139 734	1 151 305	198	200	97	1 096 193	1 127 265	6 647	6 838
	Korea	100	612 937	615 107	168	169	100	612 937	615 107	168	169	99	559 121	567 284	5 581	5 664
	Latvia	86	14 122	16 334	231	269	93	15 103	16 324	248	269	90	12 799	14 155	4 845	5 368
	Luxembourg	100	5 891	5 891	44	44	100	5 891	5 891	44	44	96	5 299	5 540	5 299	5 540
	Mexico	95	1 311 608	1 373 919	269	284	98	1 339 901	1 373 919	275	284	95	1 290 435	1 352 237	7 568	7 938
	Netherlands	63	121 527	191 966	125	201	93	178 929	191 966	184	201	85	152 346	178 985	5 345	6 269
	New Zealand	71	40 623	56 875	145	210	85	48 094	56 913	176	210	80	36 860	45 897	4 453	5 547
	Norway	95	58 824	61 809	229	241	95	58 824	61 809	229	241	91	50 163	55 277	5 456	6 016
	Poland	88	314 288	355 158	151	170	99	352 754	355 158	168	170	88	300 617	343 405	4 466	5 108
	Portugal	86	87 756	102 193	213	254	95	97 516	102 537	238	254	82	75 391	91 916	7 180	8 732
	Slovak Republic	93	50 513	54 499	272	295	99	53 908	54 562	288	295	92	45 357	49 103	6 342	6 900
	Slovenia	98	16 886	17 286	332	349	98	16 896	17 286	333	349	92	15 072	16 424	6 406	7 009
	Spain	99	404 640	409 246	199	201	100	409 246	409 246	201	201	89	356 509	399 935	6 736	7 540
	Sweden	100	93 819	94 097	202	205	100	93 819	94 097	202	205	91	82 582	91 081	5 458	6 013
	Switzerland	93	75 482	81 026	212	232	98	79 481	81 375	225	232	92	74 465	80 544	5 838	6 305
	Turkey	97	1 057 318	1 091 317	175	195	99	1 081 935	1 091 528	187	195	95	874 609	918 816	5 895	6 211
	United Kingdom	84	591 757	707 415	506	598	93	654 992	707 415	547	598	89	517 426	581 252	14 120	16 123
	United States	67	2 601 386	3 902 089	142	213	83	3 244 399	3 893 828	177	213	90	2 629 707	2 929 771	5 712	6 376
Partners	Albania	100	43 809	43 919	229	230	100	43 809	43 919	229	230	94	38 174	40 814	5 213	5 555
	Algeria	96	341 463	355 216	159	166	96	341 463	355 216	159	166	92	274 121	296 434	5 494	5 934
	Argentina	89	508 448	572 941	212	238	97	556 478	572 941	231	238	90	345 508	382 352	6 311	7 016
	Brazil	93	2 509 198	2 692 686	806	889	94	2 533 711	2 693 137	815	889	87	1 996 574	2 286 505	22 791	26 586
	B-S-J-G (China)	88	1 259 845	1 437 201	248	268	100	1 437 652	1 437 652	268	268	97	1 287 710	1 331 794	9 841	10 097
	Bulgaria	100	56 265	56 483	179	180	100	56 600	56 600	180	180	95	50 931	53 685	5 928	6 240
	Colombia	99	664 664	673 817	364	375	100	672 526	673 835	371	375	95	535 682	566 734	11 777	12 611
	Costa Rica	99	66 485	67 073	204	206	99	66 485	67 073	204	206	92	47 494	51 369	6 846	7 411
	Croatia	100	34 575	34 652	160	162	100	34 575	34 652	160	162	91	37 275	40 803	5 809	6 354
	Cyprus*	97	8 830	9 126	122	132	97	8 830	9 126	122	132	94	8 016	8 526	5 561	5 957
	Dominican Republic	99	136 669	138 187	193	195	99	136 669	138 187	193	195	94	122 620	130 700	4 731	5 026
	FYROM	100	16 426	16 472	106	107	100	16 426	16 472	106	107	95	14 999	15 802	5 324	5 617
	Georgia	97	40 552	41 595	256	267	99	41 081	41 566	262	267	94	35 567	37 873	5 316	5 689
	Hong Kong (China)	75	45 603	60 716	115	153	90	54 795	60 715	138	153	93	48 222	51 806	5 359	5 747
	Indonesia	98	3 126 468	3 176 076	232	236	100	3 176 076	3 176 076	236	236	98	3 015 844	3 092 773	6 513	6 694
	Jordan	100	119 024	119 024	250	250	100	119 024	119 024	250	250	97	105 868	108 669	7 267	7 462
	Kazakhstan	100	202 701	202 701	232	232	100	202 701	202 701	232	232	97	187 683	192 921	7 841	8 059
	Kosovo	100	26 924	26 924	224	224	100	26 924	26 924	224	224	99	22 016	22 333	4 826	4 896
	Lebanon	67	40 542	60 882	208	308	87	53 091	60 797	270	308	95	36 052	38 143	4 546	4 788
	Lithuania	99	31 386	31 588	309	311	100	31 543	31 588	310	311	91	27 070	29 889	6 523	7 202
	Macao (China)	100	4 414	4 414	45	45	100	4 414	4 414	45	45	99	4 476	4 507	4 476	4 507
	Malaysia	51	229 340	446 237	147	230	98	437 424	446 100	224	230	97	393 785	407 396	8 843	9 097
	Malta	100	4 341	4 343	59	61	100	4 341	4 343	59	61	85	3 634	4 294	3 634	4 294
	Moldova	100	30 145	30 145	229	229	100	30 145	30 145	229	229	98	28 754	29 341	5 325	5 436
	Montenegro	100	7 301	7 312	64	65	100	7 301	7 312	64	65	94	6 346	6 766	5 665	6 043
	Peru	100	468 406	470 651	280	282	100	469 662	470 651	281	282	99	426 205	430 959	6 971	7 054
	Qatar	99	13 333	13 470	166	168	99	13 333	13 470	166	168	94	12 061	12 819	12 061	12 819
	Romania	99	171 553	172 652	181	182	100	172 495	172 495	182	182	99	162 918	164 216	4 876	4 910
	Russia	99	1 181 937	1 189 441	209	210	99	1 181 937	1 189 441	209	210	97	1 072 914	1 108 068	6 021	6 215
	Singapore	97	45 299	46 620	175	179	98	45 553	46 620	176	179	93	42 241	45 259	6 105	6 555
	Chinese Taipei	100	286 778	286 778	214	214	100	286 778	286 778	214	214	98	246 408	251 424	7 708	7 871
	Thailand	99	739 772	751 010	269	273	100	751 010	751 010	273	273	97	614 996	634 795	8 249	8 491
	Trinidad and Tobago	92	15 904	17 371	141	163	92	15 904	17 371	141	163	79	9 674	12 188	4 587	5 745
	Tunisia	99	121 751	122 767	162	165	99	121 838	122 792	163	165	86	97 337	112 665	5 340	6 175
	United Arab Emirates	99	49 310	50 060	473	477	99	49 310	50 060	473	477	95	43 774	46 263	14 167	15 014
	Uruguay	98	42 986	43 737	217	221	99	43 442	43 737	219	221	86	32 762	38 023	6 059	7 026
	Viet Nam	100	996 757	996 757	188	188	100	996 757	996 757	188	188	100	871 353	874 859	5 826	5 849

* See note at the beginning of this Annex.

StatLink http://dx.doi.org/10.1787/888933433129

Grade levels

Students assessed in PISA 2015 are at various grade levels. The percentage of students at each grade level is presented by country in Table A2.4a and by gender within each country 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)	11.2	(0.3)	74.6	(0.4)	14.0	(0.4)	0.1	(0.0)
	Austria	0.0	(0.0)	2.0	(0.6)	20.8	(0.9)	71.2	(1.0)	5.9	(0.3)	0.0	(0.0)
	Belgium	0.6	(0.1)	6.4	(0.5)	30.7	(0.7)	61.0	(0.9)	1.3	(0.1)	0.0	(0.0)
	Canada	0.1	(0.0)	0.7	(0.1)	10.8	(0.5)	87.6	(0.6)	0.8	(0.1)	0.0	(0.0)
	Chile	1.7	(0.3)	4.1	(0.6)	24.0	(0.7)	68.1	(1.0)	2.1	(0.2)	0.0	(0.0)
	Czech Republic	0.5	(0.1)	3.9	(0.3)	49.4	(1.2)	46.2	(1.2)	0.0	(0.0)	0.0	c
	Denmark	0.2	(0.1)	16.4	(0.6)	81.9	(0.7)	1.4	(0.5)	0.0	c	0.0	c
	Estonia	0.8	(0.2)	21.3	(0.6)	76.6	(0.6)	1.3	(0.3)	0.0	c	0.0	(0.0)
	Finland	0.5	(0.1)	13.6	(0.4)	85.7	(0.4)	0.0	(0.0)	0.2	(0.1)	0.0	c
	France	0.0	(0.0)	1.0	(0.2)	23.1	(0.6)	72.5	(0.7)	3.2	(0.2)	0.1	(0.1)
	Germany	0.5	(0.1)	7.7	(0.4)	47.3	(0.8)	43.1	(0.8)	1.5	(0.5)	0.0	(0.0)
	Greece	0.2	(0.1)	0.7	(0.2)	3.8	(0.8)	95.3	(0.9)	0.0	c	0.0	c
	Hungary	1.7	(0.3)	8.5	(0.5)	75.8	(0.7)	14.0	(0.5)	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)	1.8	(0.2)	60.6	(0.7)	26.5	(1.1)	11.1	(0.9)	0.0	c
	Israel	0.0	c	0.1	(0.0)	16.4	(0.9)	82.7	(0.9)	0.9	(0.3)	0.0	c
	Italy	0.1	(0.0)	1.0	(0.2)	15.2	(0.6)	77.2	(0.7)	6.6	(0.3)	0.0	c
	Japan	0.0	c	0.0	c	0.0	c	100.0	(0.0)	0.0	c	0.0	c
	Korea	0.0	c	0.0	c	9.1	(0.8)	90.4	(0.8)	0.5	(0.1)	0.0	c
	Latvia	0.9	(0.2)	11.7	(0.5)	84.4	(0.6)	2.9	(0.3)	0.0	(0.0)	0.0	c
	Luxembourg	0.3	(0.1)	7.9	(0.1)	50.9	(0.1)	40.3	(0.1)	0.6	(0.0)	0.0	c
	Mexico	2.3	(0.3)	4.8	(0.4)	31.9	(1.4)	60.3	(1.6)	0.5	(0.1)	0.2	(0.0)
	Netherlands	0.1	(0.0)	2.8	(0.3)	41.6	(0.6)	54.8	(0.6)	0.8	(0.2)	0.0	(0.0)
	New Zealand	0.0	c	0.0	c	0.0	(0.0)	6.2	(0.3)	88.8	(0.5)	5.0	(0.5)
	Norway	0.0	c	0.0	c	0.6	(0.1)	99.3	(0.2)	0.1	(0.1)	0.0	c
	Poland	0.6	(0.1)	4.9	(0.3)	93.8	(0.4)	0.6	(0.2)	0.0	c	0.0	c
	Portugal	3.2	(0.3)	8.4	(0.5)	22.9	(0.9)	65.1	(1.2)	0.4	(0.1)	0.0	c
	Slovak Republic	2.2	(0.4)	4.6	(0.4)	42.6	(1.3)	50.6	(1.2)	0.1	(0.0)	0.0	c
	Slovenia	0.0	c	0.3	(0.1)	4.8	(0.3)	94.6	(0.4)	0.3	(0.1)	0.0	c
	Spain	0.1	(0.0)	8.6	(0.5)	23.4	(0.6)	67.9	(0.9)	0.1	(0.1)	0.0	c
	Sweden	0.1	(0.1)	3.1	(0.4)	94.9	(0.8)	1.8	(0.7)	0.1	(0.1)	0.0	c
	Switzerland	0.5	(0.1)	11.8	(0.7)	61.3	(1.2)	25.9	(1.3)	0.5	(0.1)	0.0	(0.0)
	Turkey	0.6	(0.1)	2.6	(0.4)	20.7	(1.0)	72.9	(1.2)	3.0	(0.3)	0.1	(0.0)
	United Kingdom	0.0	c	0.0	c	0.0	c	1.6	(0.3)	97.4	(0.4)	1.0	(0.3)
	United States	0.0	(0.0)	0.5	(0.3)	9.6	(0.7)	72.4	(0.9)	17.3	(0.6)	0.1	(0.0)
Partners	Albania	0.2	(0.1)	1.0	(0.2)	35.8	(2.3)	61.7	(2.3)	1.2	(0.7)	0.0	(0.0)
	Algeria	18.8	(1.0)	23.5	(1.1)	35.1	(1.5)	19.4	(2.1)	3.2	(0.7)	0.0	c
	Brazil	3.5	(0.2)	6.4	(0.4)	12.5	(0.5)	35.9	(0.9)	39.2	(0.8)	2.5	(0.2)
	B-S-J-G (China)	1.1	(0.2)	9.2	(0.7)	52.7	(1.7)	34.6	(2.0)	2.2	(0.5)	0.1	(0.0)
	Bulgaria	0.5	(0.2)	3.0	(0.6)	92.2	(0.8)	4.3	(0.4)	0.0	c	0.0	c
	Colombia	5.3	(0.4)	12.3	(0.6)	22.7	(0.6)	40.2	(0.7)	19.5	(0.6)	0.0	c
	Costa Rica	6.2	(0.7)	14.0	(0.7)	33.0	(1.2)	46.5	(1.6)	0.2	(0.1)	0.1	(0.1)
	Croatia	0.0	c	0.2	(0.2)	79.2	(0.5)	20.6	(0.4)	0.0	c	0.0	c
	Cyprus*	0.0	c	0.3	(0.0)	5.8	(0.1)	93.1	(0.1)	0.7	(0.1)	0.0	c
	Dominican Republic	7.1	(0.8)	13.8	(1.2)	20.6	(0.8)	41.9	(1.1)	14.2	(0.7)	2.4	(0.3)
	FYROM	0.1	(0.1)	0.1	(0.1)	70.2	(0.2)	29.7	(0.2)	0.0	c	0.0	c
	Georgia	0.1	(0.0)	0.8	(0.2)	22.0	(0.8)	76.0	(0.9)	1.1	(0.3)	0.0	c
	Hong Kong (China)	1.1	(0.1)	5.6	(0.4)	26.0	(0.7)	66.7	(0.7)	0.6	(0.5)	0.0	c
	Indonesia	2.1	(0.3)	8.1	(0.7)	42.1	(1.5)	45.5	(1.6)	2.3	(0.4)	0.0	(0.0)
	Jordan	0.2	(0.1)	0.6	(0.1)	6.6	(0.4)	92.6	(0.4)	0.0	c	0.0	c
	Kosovo	0.0	(0.1)	0.6	(0.1)	24.9	(0.8)	72.4	(0.9)	2.1	(0.2)	0.0	c
	Lebanon	3.7	(0.5)	8.3	(0.8)	16.6	(1.1)	62.3	(1.4)	9.0	(0.8)	0.1	(0.1)
	Lithuania	0.1	(0.0)	2.6	(0.2)	86.3	(0.4)	11.0	(0.4)	0.0	(0.0)	0.0	c
	Macao (China)	2.9	(0.1)	12.2	(0.2)	29.7	(0.2)	54.5	(0.1)	0.6	(0.1)	0.0	c
	Malta	0.0	c	0.0	c	0.3	(0.1)	6.1	(0.2)	93.6	(0.1)	0.1	(0.0)
	Moldova	0.2	(0.1)	7.6	(0.5)	84.5	(0.8)	7.5	(0.8)	0.0	(0.0)	0.0	c
	Montenegro	0.0	c	0.0	c	83.7	(0.1)	16.3	(0.1)	0.0	c	0.0	c
	Peru	2.5	(0.3)	6.6	(0.4)	15.9	(0.5)	50.2	(0.8)	24.8	(0.8)	0.0	c
	Qatar	0.9	(0.1)	3.5	(0.1)	16.3	(0.1)	60.7	(0.1)	18.0	(0.1)	0.6	(0.0)
	Romania	1.4	(0.3)	8.9	(0.5)	74.8	(0.9)	14.9	(0.7)	0.0	c	0.0	c
	Russia	0.2	(0.1)	6.6	(0.3)	79.7	(1.5)	13.4	(1.5)	0.1	(0.0)	0.0	c
	Singapore	0.0	(0.0)	1.9	(0.3)	7.9	(0.8)	90.0	(1.0)	0.1	(0.0)	0.1	(0.0)
	Chinese Taipei	0.0	c	0.0	c	35.4	(0.7)	64.6	(0.7)	0.0	c	0.0	c
	Thailand	0.2	(0.1)	0.6	(0.2)	23.8	(1.0)	72.9	(1.0)	2.4	(0.4)	0.0	c
	Trinidad and Tobago	3.3	(0.2)	10.8	(0.3)	27.3	(0.3)	56.5	(0.3)	2.2	(0.2)	0.0	c
	Tunisia	4.3	(0.3)	10.6	(0.8)	19.6	(1.3)	60.9	(1.7)	4.6	(0.4)	0.0	c
	United Arab Emirates	0.6	(0.1)	2.5	(0.3)	10.6	(0.7)	53.4	(0.8)	31.4	(0.8)	1.5	(0.1)
	Uruguay	7.5	(0.6)	9.7	(0.5)	20.7	(0.7)	61.3	(1.2)	0.8	(0.1)	0.0	c
	Viet Nam	0.3	(0.1)	1.7	(0.4)	7.7	(1.8)	90.4	(2.2)	0.0	(0.0)	0.0	c
	Argentina**	1.6	(0.4)	9.7	(0.8)	27.4	(1.2)	58.5	(1.6)	2.8	(0.3)	0.0	c
	Kazakhstan**	0.1	(0.1)	2.7	(0.3)	60.4	(1.7)	36.2	(1.8)	0.6	(0.1)	0.0	c
	Malaysia**	0.0	c	0.0	c	3.2	(0.6)	96.4	(0.7)	0.4	(0.3)	0.0	c

* See note at the beginning of this Annex.
** Coverage is too small to ensure comparability (see Annex A4).
StatLink http://dx.doi.org/10.1787/888933433129

[Part 1/1]

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

		Boys												Girls											
		7th grade		8th grade		9th grade		10th grade		11th grade		12th grade and above		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.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.	%	S.E.
OECD	Australia	0.0	(0.0)	0.2	(0.1)	13.2	(0.4)	73.5	(0.5)	13.1	(0.5)	0.0	(0.0)	0.0	(0.0)	0.1	(0.0)	9.2	(0.3)	75.7	(0.5)	14.9	(0.6)	0.1	(0.1)
	Austria	0.1	(0.1)	2.0	(0.4)	21.6	(1.2)	71.1	(1.2)	5.2	(0.4)	0.0	(0.0)	0.0	c	2.0	(0.9)	20.0	(1.0)	71.4	(1.3)	6.6	(0.4)	0.0	(0.0)
	Belgium	0.7	(0.1)	6.7	(0.5)	33.6	(1.0)	57.9	(1.1)	1.2	(0.2)	0.0	c	0.6	(0.1)	6.2	(0.5)	27.7	(0.8)	64.2	(1.1)	1.3	(0.1)	0.0	(0.0)
	Canada	0.1	(0.1)	1.0	(0.2)	11.7	(0.6)	86.5	(0.6)	0.7	(0.1)	0.0	(0.0)	0.1	(0.0)	0.4	(0.1)	9.9	(0.6)	88.8	(0.6)	0.8	(0.1)	0.0	(0.0)
	Chile	2.2	(0.5)	4.8	(0.8)	26.4	(0.9)	64.8	(1.3)	1.8	(0.2)	0.1	(0.1)	1.2	(0.4)	3.5	(0.7)	21.5	(0.8)	71.4	(1.1)	2.4	(0.3)	0.0	c
	Czech Republic	0.6	(0.2)	5.5	(0.5)	52.3	(1.5)	41.5	(1.6)	0.0	(0.0)	0.0	c	0.4	(0.2)	2.2	(0.3)	46.2	(1.5)	51.2	(1.6)	0.0	c	0.0	c
	Denmark	0.3	(0.1)	21.9	(0.9)	76.6	(1.0)	1.2	(0.5)	0.0	c	0.0	c	0.1	(0.1)	10.8	(0.5)	87.3	(0.7)	1.7	(0.6)	0.0	c	0.0	c
	Estonia	1.3	(0.3)	23.7	(0.9)	74.2	(0.8)	0.8	(0.3)	0.0	c	0.0	(0.0)	0.2	(0.1)	18.8	(0.8)	79.1	(0.8)	1.9	(0.4)	0.0	c	0.0	c
	Finland	0.4	(0.1)	15.5	(0.6)	83.9	(0.6)	0.0	(0.0)	0.2	(0.1)	0.0	c	0.5	(0.1)	11.5	(0.5)	87.7	(0.5)	0.0	c	0.3	(0.2)	0.0	c
	France	0.0	c	1.0	(0.2)	26.1	(0.9)	69.6	(1.0)	3.1	(0.3)	0.2	(0.1)	0.1	(0.1)	1.0	(0.2)	20.1	(0.6)	75.4	(0.8)	3.3	(0.3)	0.1	(0.0)
	Germany	0.7	(0.2)	9.0	(0.5)	50.1	(1.0)	38.8	(1.0)	1.4	(0.4)	0.0	(0.0)	0.3	(0.1)	6.3	(0.6)	44.3	(0.9)	47.5	(1.0)	1.6	(0.6)	0.0	c
	Greece	0.4	(0.2)	1.1	(0.3)	4.7	(1.0)	93.8	(1.2)	0.0	c	0.0	c	0.1	(0.1)	0.2	(0.1)	2.8	(0.8)	96.9	(0.8)	0.0	c	0.0	c
	Hungary	1.8	(0.4)	10.1	(0.6)	75.6	(0.9)	12.5	(0.6)	0.0	c	0.0	c	1.6	(0.4)	6.9	(0.8)	76.0	(0.9)	15.5	(0.7)	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	0.0	c	0.0	c	0.0	c	100.0	c	0.0	c	0.0	c
	Ireland	0.0	c	2.2	(0.3)	62.8	(0.9)	24.1	(1.2)	10.9	(1.0)	0.0	c	0.0	(0.0)	1.4	(0.2)	58.2	(0.9)	29.0	(1.4)	11.3	(1.1)	0.0	c
	Israel	0.0	c	0.1	(0.1)	18.0	(1.2)	80.9	(1.3)	1.1	(0.6)	0.0	c	0.0	c	0.1	(0.0)	14.9	(0.8)	84.4	(0.8)	0.7	(0.1)	0.0	c
	Italy	0.2	(0.1)	1.3	(0.3)	18.1	(0.8)	75.0	(0.9)	5.4	(0.4)	0.0	c	0.1	(0.0)	0.7	(0.2)	12.2	(0.8)	79.3	(1.0)	7.7	(0.5)	0.0	c
	Japan	0.0	c	0.0	c	0.0	c	100.0	c	0.0	c	0.0	c	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	10.1	(1.4)	89.4	(1.4)	0.5	(0.1)	0.0	c	0.0	c	0.0	c	8.0	(0.8)	91.5	(0.8)	0.5	(0.1)	0.0	c
	Latvia	1.5	(0.4)	14.7	(0.8)	81.8	(0.9)	1.9	(0.3)	0.0	(0.0)	0.0	c	0.4	(0.2)	8.7	(0.7)	87.0	(0.7)	3.9	(0.4)	0.0	c	0.0	c
	Luxembourg	0.2	(0.1)	9.4	(0.2)	52.4	(0.3)	37.3	(0.2)	0.7	(0.1)	0.0	c	0.3	(0.1)	6.4	(0.2)	49.4	(0.2)	43.3	(0.2)	0.6	(0.1)	0.0	c
	Mexico	3.1	(0.5)	5.9	(0.6)	32.2	(1.5)	58.0	(1.6)	0.6	(0.2)	0.2	(0.0)	1.5	(0.3)	3.7	(0.4)	31.6	(1.7)	62.5	(1.7)	0.4	(0.1)	0.2	(0.1)
	Netherlands	0.0	(0.0)	3.8	(0.4)	45.3	(0.8)	50.2	(0.8)	0.8	(0.3)	0.0	c	0.1	(0.0)	1.9	(0.3)	38.0	(0.7)	59.3	(0.7)	0.7	(0.2)	0.0	(0.0)
	New Zealand	0.0	c	0.0	c	0.0	c	6.9	(0.5)	88.6	(0.8)	4.5	(0.5)	0.0	c	0.0	c	0.0	(0.0)	5.4	(0.4)	89.1	(0.6)	5.5	(0.6)
	Norway	0.0	c	0.0	c	0.8	(0.2)	99.1	(0.2)	0.1	(0.1)	0.0	c	0.0	c	0.0	c	0.3	(0.1)	99.6	(0.1)	0.1	(0.1)	0.0	c
	Poland	0.9	(0.2)	6.8	(0.5)	92.1	(0.6)	0.2	(0.2)	0.0	c	0.0	c	0.4	(0.1)	3.0	(0.3)	95.6	(0.5)	1.1	(0.3)	0.0	c	0.0	c
	Portugal	4.2	(0.4)	10.5	(0.7)	25.4	(1.0)	59.6	(1.4)	0.3	(0.1)	0.0	c	2.1	(0.4)	6.4	(0.5)	20.5	(0.9)	70.5	(1.2)	0.5	(0.1)	0.0	c
	Slovak Republic	2.4	(0.4)	4.8	(0.5)	43.5	(1.6)	49.4	(1.8)	0.0	c	0.0	c	1.9	(0.5)	4.3	(0.6)	41.7	(1.8)	51.9	(1.8)	0.1	(0.1)	0.0	c
	Slovenia	0.0	c	0.5	(0.2)	5.4	(0.7)	93.9	(0.7)	0.2	(0.1)	0.0	c	0.0	c	0.2	(0.1)	4.1	(0.6)	95.3	(0.6)	0.4	(0.2)	0.0	c
	Spain	0.1	(0.1)	10.7	(0.7)	25.4	(0.8)	63.7	(1.1)	0.1	(0.1)	0.0	c	0.0	c	6.5	(0.5)	21.3	(0.8)	72.1	(1.0)	0.1	(0.1)	0.0	c
	Sweden	0.1	(0.1)	3.5	(0.5)	95.0	(0.9)	1.4	(0.7)	0.1	(0.1)	0.0	c	0.2	(0.1)	2.6	(0.4)	94.9	(1.0)	2.3	(0.9)	0.1	(0.1)	0.0	c
	Switzerland	0.7	(0.2)	13.4	(0.8)	60.7	(1.1)	24.7	(1.2)	0.5	(0.1)	0.0	c	0.3	(0.1)	10.1	(0.8)	62.0	(1.7)	27.2	(1.9)	0.5	(0.2)	0.0	(0.0)
	Turkey	0.8	(0.3)	3.1	(0.6)	25.4	(1.2)	68.4	(1.6)	2.2	(0.4)	0.1	(0.1)	0.4	(0.2)	2.1	(0.4)	16.1	(1.1)	77.5	(1.3)	3.8	(0.4)	0.1	(0.0)
	United Kingdom	0.0	c	0.0	c	0.0	c	1.9	(0.5)	97.3	(0.6)	0.9	(0.3)	0.0	c	0.0	c	0.0	c	1.4	(0.2)	97.5	(0.3)	1.1	(0.3)
	United States	0.0	c	0.5	(0.4)	11.6	(0.8)	72.4	(1.0)	15.3	(0.7)	0.2	(0.1)	0.1	(0.1)	0.5	(0.2)	7.6	(0.6)	72.4	(0.9)	19.4	(0.7)	0.1	(0.0)
Partners	Albania	0.2	(0.2)	0.9	(0.2)	41.2	(2.7)	56.3	(2.6)	1.3	(0.9)	0.0	(0.0)	0.1	(0.1)	1.1	(0.3)	30.4	(2.1)	67.1	(2.2)	1.2	(0.5)	0.1	(0.0)
	Algeria	24.4	(1.3)	25.7	(1.2)	32.6	(1.5)	14.7	(1.9)	2.6	(0.7)	0.0	c	12.6	(1.1)	21.0	(1.2)	37.9	(2.0)	24.6	(2.5)	3.9	(0.8)	0.0	c
	Brazil	4.6	(0.3)	7.8	(0.6)	13.9	(0.6)	36.5	(1.0)	35.3	(0.9)	1.8	(0.2)	2.4	(0.2)	5.0	(0.4)	11.1	(0.6)	35.3	(0.9)	43.0	(0.9)	3.1	(0.2)
	B-S-J-G (China)	1.2	(0.2)	9.9	(0.7)	55.4	(1.7)	31.6	(1.9)	1.9	(0.5)	0.1	(0.0)	1.1	(0.2)	8.4	(0.8)	49.6	(1.8)	38.1	(2.2)	2.6	(0.5)	0.1	(0.1)
	Bulgaria	0.6	(0.2)	4.1	(0.8)	91.8	(1.0)	3.5	(0.4)	0.0	c	0.0	c	0.4	(0.2)	1.8	(0.4)	92.7	(0.7)	5.2	(0.4)	0.0	c	0.0	c
	Colombia	7.2	(0.6)	14.3	(0.8)	25.2	(0.8)	37.1	(0.9)	16.2	(0.8)	0.0	c	3.6	(0.4)	10.5	(0.7)	20.5	(0.9)	42.9	(1.0)	22.5	(0.8)	0.0	c
	Costa Rica	7.8	(0.8)	16.7	(0.8)	34.3	(1.2)	41.2	(1.5)	0.1	(0.0)	0.0	c	4.7	(0.7)	11.4	(0.7)	31.8	(1.4)	51.6	(1.8)	0.3	(0.1)	0.2	(0.1)
	Croatia	0.0	c	0.2	(0.1)	80.5	(0.5)	19.4	(0.5)	0.0	c	0.0	c	0.0	c	0.3	(0.2)	78.0	(0.7)	21.7	(0.7)	0.0	c	0.0	c
	Cyprus*	0.0	c	0.3	(0.1)	6.6	(0.2)	92.4	(0.2)	0.6	(0.1)	0.0	c	0.0	c	0.3	(0.1)	5.1	(0.2)	93.8	(0.2)	0.8	(0.1)	0.0	c
	Dominican Republic	10.3	(1.1)	16.4	(1.5)	23.3	(1.2)	37.2	(1.4)	11.1	(0.8)	1.7	(0.3)	4.0	(0.6)	11.2	(1.1)	18.1	(0.8)	46.5	(1.1)	17.2	(0.8)	3.0	(0.3)
	FYROM	0.2	(0.2)	0.2	(0.2)	70.9	(0.3)	28.8	(0.2)	0.0	c	0.0	c	0.0	c	0.0	c	69.4	(0.3)	30.6	(0.3)	0.0	c	0.0	c
	Georgia	0.1	(0.0)	0.9	(0.2)	23.0	(1.0)	75.2	(1.0)	0.8	(0.2)	0.0	c	0.1	(0.1)	0.7	(0.2)	20.9	(0.9)	76.8	(1.0)	1.5	(0.4)	0.0	c
	Hong Kong (China)	1.3	(0.2)	6.4	(0.5)	28.5	(0.8)	63.3	(0.9)	0.5	(0.4)	0.0	c	1.0	(0.2)	4.7	(0.4)	23.5	(0.8)	70.2	(0.9)	0.6	(0.6)	0.0	c
	Indonesia	2.5	(0.4)	8.9	(0.9)	44.3	(1.9)	42.1	(2.0)	2.1	(0.4)	0.0	(0.0)	1.7	(0.3)	7.2	(1.0)	39.8	(1.9)	48.9	(2.1)	2.4	(0.4)	0.0	c
	Jordan	0.1	(0.1)	0.5	(0.1)	6.6	(0.7)	92.9	(0.7)	0.0	c	0.0	c	0.2	(0.1)	0.7	(0.1)	6.6	(0.6)	92.4	(0.6)	0.0	c	0.0	c
	Kosovo	0.1	(0.1)	0.5	(0.1)	26.4	(0.9)	71.5	(1.0)	1.6	(0.3)	0.0	c	0.0	c	0.7	(0.2)	23.5	(1.0)	73.3	(1.0)	2.5	(0.3)	0.0	c
	Lebanon	4.0	(0.6)	8.2	(0.9)	17.2	(1.4)	63.5	(1.7)	6.9	(0.7)	0.2	(0.1)	3.4	(0.6)	8.3	(1.0)	16.1	(1.2)	61.2	(1.8)	10.8	(1.2)	0.1	(0.1)
	Lithuania	0.2	(0.1)	3.5	(0.3)	87.4	(0.6)	8.8	(0.5)	0.0	(0.0)	0.0	c	0.0	(0.0)	1.7	(0.2)	85.1	(0.7)	13.1	(0.6)	0.0	(0.0)	0.0	c
	Macao (China)	4.3	(0.2)	16.4	(0.3)	30.8	(0.2)	48.2	(0.2)	0.4	(0.1)	0.0	c	1.6	(0.2)	8.0	(0.2)	28.7	(0.3)	60.8	(0.3)	0.9	(0.2)	0.0	c
	Malta	0.0	c	0.0	c	0.5	(0.1)	6.8	(0.3)	92.7	(0.2)	0.0	c	0.0	c	0.0	c	0.1	(0.0)	5.4	(0.2)	94.4	(0.2)	0.1	(0.1)
	Moldova	0.3	(0.1)	8.2	(0.7)	86.3	(0.9)	5.0	(0.9)	0.1	(0.1)	0.0	c	0.2	(0.1)	7.0	(0.6)	82.8	(1.2)	10.1	(1.2)	0.0	c	0.0	c
	Montenegro	0.0	c	0.0	c	85.2	(0.2)	14.8	(0.2)	0.0	c	0.0	c	0.0	c	0.0	c	82.2	(0.2)	17.8	(0.2)	0.0	c	0.0	c
	Peru	3.0	(0.5)	7.5	(0.5)	17.9	(0.7)	48.7	(0.9)	22.9	(1.0)	0.0	c	1.9	(0.3)	5.6	(0.5)	14.0	(0.6)	51.7	(1.0)	26.8	(0.9)	0.0	c
	Qatar	0.8	(0.1)	3.6	(0.1)	18.0	(0.2)	59.3	(0.2)	17.6	(0.2)	0.6	(0.1)	1.0	(0.1)	3.4	(0.1)	14.5	(0.1)	62.1	(0.2)	18.4	(0.2)	0.6	(0.1)
	Romania	1.7	(0.4)	10.7	(0.8)	74.3	(1.0)	13.3	(0.7)	0.0	c	0.0	c	1.1	(0.4)	7.2	(0.8)	75.3	(1.1)	16.4	(0.8)	0.0	c	0.0	c
	Russia	0.2	(0.1)	7.2	(0.5)	80.1	(1.7)	12.4	(1.7)	0.0	(0.0)	0.0	c	0.1	(0.1)	6.0	(0.4)	79.3	(1.5)	14.4	(1.6)	0.1	(0.1)	0.0	c
	Singapore	0.1	(0.0)	1.8	(0.3)	8.9	(0.9)	89.1	(1.1)	0.1	(0.1)	0.0	(0.0)	0.0	(0.0)	2.0	(0.4)	6.9	(0.8)	90.8	(1.1)	0.2	(0.1)	0.1	(0.0)
	Chinese Taipei	0.0	c	0.0	c	36.5	(1.3)	63.5	(1.3)	0.0	c	0.0	c	0.0	c	0.0	c	34.3	(1.3)	65.7	(1.3)	0.0	c	0.0	c
	Thailand	0.2	(0.1)	0.8	(0.3)	25.4	(1.2)	71.4	(1.2)	2.3	(0.4)	0.0	c	0.3	(0.1)	0.5	(0.2)	22.5	(1.3)	74.1	(1.3)	2.6	(0.4)	0.0	c
	Trinidad and Tobago	3.7	(0.3)	14.2	(0.5)	30.8	(0.5)	48.9	(0.5)	2.4	(0.2)	0.0	c	2.8	(0.2)	7.5	(0.4)	23.8	(0.4)	63.9	(0.5)	2.0	(0.3)	0.0	c
	Tunisia	5.9	(0.5)	13.8	(1.0)	22.0	(1.4)	54.0	(1.9)	4.3	(0.5)	0.0	c	3.0	(0.3)	7.8	(0.7)	17.5	(1.4)	67.0	(1.8)	4.8	(0.5)	0.0	c
	United Arab Emirates	0.7	(0.1)	2.9	(0.4)	11.4	(1.1)	54.0	(1.3)	29.6	(1.0)	1.4	(0.2)	0.4	(0.1)	2.2	(0.5)	9.9	(0.9)	52.8	(0.9)	33.1	(1.1)	1.6	(0.2)
	Uruguay	9.2	(0.8)	11.2	(0.7)	22.5	(0.9)	56.5	(1.5)	0.5	(0.1)	0.0	c	6.0	(0.7)	8.3	(0.6)	19.0	(0.8)	65.6	(1.1)	1.1	(0.2)	0.0	c
	Viet Nam	0.5	(0.2)	2.3	(0.6)	11.1	(2.6)	86.1	(3.2)	0.0	c	0.0	c	0.1	(0.0)	1.1	(0.4)	4.6	(1.2)	94.2	(1.4)	0.0	(0.0)	0.0	c
	Argentina**	2.3	(0.6)	11.5	(0.9)	27.8	(1.3)	56.0	(1.8)	2.4	(0.3)	0.0	c	1.0	(0.3)	8.1	(0.9)	26.9	(1.4)	60.8	(1.7)	3.2	(0.3)	0.0	c
	Kazakhstan**	0.1	(0.1)	3.1	(0.4)	62.8	(2.3)	33.5	(2.4)	0.5	(0.1)	0.0	c	0.1	(0.1)	2.3	(0.3)	57.8	(1.7)	39.0	(1.8)	0.7	(0.1)	0.0	c
	Malaysia**	0.0	c	0.0	c	4.2	(0.8)	95.4	(0.9)	0.4	(0.3)	0.0	c	0.0	c	0.0	c	2.3	(0.5)	97.2	(0.6)	0.4	(0.4)	0.0	c

* See note at the beginning of this Annex.
** Coverage is too small to ensure comparability (see Annex A4).
StatLink http://dx.doi.org/10.1787/888933433129

Reference

OECD (forthcoming), *PISA 2015 Technical Report*, PISA, OECD Publishing, Paris.

ANNEX A3
TECHNICAL NOTES ON ANALYSES IN PISA 2015 RESULTS

Methods and definitions

Odds ratio

The odds ratio is a measure of the relative likelihood of a particular outcome across two groups. The odds ratio for observing the outcome when an antecedent is present is simply

$$OR = \frac{(p_{11}/p_{12})}{(p_{21}/p_{22})}$$

where p_{11}/p_{12} represents the "odds" of observing the outcome when the antecedent is present, and p_{21}/p_{22} represents the "odds" of observing the outcome when the antecedent is not present.

Logistic regression can be used to estimate the log ratio: the exponentiated logit coefficient for a binary variable is equivalent to the odds ratio. A "generalised" odds ratio, after accounting for other differences across groups, can be estimated by introducing control variables in the logistic regression.

Statistics based on multilevel models

Statistics based on multilevel 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 science, reading or mathematics performance, the estimation uses ten plausible values for each student's performance on the mathematics scale. Models were estimated using the Stata ® (version 14.1) "mixed" module. The three-level regression models are estimated with HLM® (version 6.06) using only five plausible values of science performance.

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 amount to the sample size within each school. Between-school weights correspond to the sum of final student weights (W_FSTUWT) within each school. The definition of between-school weights is the same as in PISA 2012 initial reports. For the three-level regression models, the sum of the weights is the same across education systems so that each education system contributes equally to the results.

The index of inclusion is based on the intraclass correlation and is estimated as:

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

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

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 2015 Technical Report* (OECD, forthcoming) and Annex A2 provide an overview of how schools are defined. In Slovenia, the primary sampling unit is defined as a group of students who follow the same study programme within a school (an education track within a school). So in this case, the between-school variation is actually the between-track variation. The use of stratification variables in the selection of schools may also affect the estimate of the between-school variation, 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

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. For the three-level regression models, five imputed values were computed for each missing value using the predictive mean matching method in SAS® PROC MI. Five plausible values of science performance were then analysed by the HLM® software using one of the five imputed data sets.

Diversity index of grade levels

The diversity index of grade levels is based on the Herfindahl index and can be interpreted as the probability (in %) that two students selected at random are enrolled in different grades. It is defined as:

$$D = 100 - ((\sum_{g=1}^{G} p_g^2) * 100))$$

where p_g is the proportion of students enrolled in grade level g.

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 if the probability of reporting a difference when there is actually no such difference in corresponding population values is lower than 5%. 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.

Differences between subgroup means

Differences between groups of students (e.g. students who have skipped a day of school and students who have not skipped a day of school) or categories of schools (e.g. advantaged and disadvantaged schools) were tested for statistical significance. The definitions of the subgroups can, in general, be found in the tables and the text accompanying the analysis. Socio-economically (dis)advantaged schools, for instance, are defined as schools in the (bottom) top quarter of the distribution of the average PISA index of economic, social and cultural status (ESCS) across schools within each country/economy. All differences marked in bold in the tables presented in Annex B of this report are statistically significant at the 95% level.

Change in the performance per unit of an index

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

Odds ratio

Figures in bold in the data tables presented in Annex B of this report indicate that the relative risk/odds ratio 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/odds-ratio statistic is assumed to follow a log-normal distribution, rather than a normal distribution, under the null hypothesis.

Multilevel models

The standard errors of multilevel models 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.

Multiple imputation

The standard errors take into account the between-imputation variance. The standard errors of the results therefore consist of sampling variance, cognitive test measurement variance and error due to the imputation of missing values.

Reference

Gorard, S. and **C. Taylor** (2002), "What is segregation ? A comparison of measures in terms of 'strong' and 'weak' compositional invariance", *Sociology*, Vol.36/4, pp. 875-895, http://dx.doi.org/ 10.1177/003803850203600405.

ANNEX A4
QUALITY ASSURANCE

Quality assurance procedures were implemented in all parts of PISA 2015, as was done for all previous PISA surveys. The PISA 2015 Technical Standards (www.oecd.org/pisa/) specify the way in which PISA must be implemented in each country, economy and adjudicated region. International contractors monitor the implementation in each of these and adjudicate on their adherence to the standards.

The consistent quality and linguistic equivalence of the PISA 2015 assessment instruments were facilitated by assessing the ease with which the original English version could be translated. Two source versions of the assessment instruments, in English and French were prepared (except for the financial literacy assessment and the operational manuals, which were provided only in English) in order for countries to conduct a double translation design, i.e. two independent translations from the source language(s), and reconciliation by a third person. Detailed instructions for the localisation (adaptation, translation and validation) of the instruments for the field trial and for their review for the main survey, and translation/adaptation guidelines were supplied. An independent team of expert verifiers, appointed and trained by the PISA Consortium, verified each national version against the English and/or French source versions. These translators' mother tongue was the language of instruction in the country concerned, and the translators were knowledgeable about education systems. For further information on PISA translation procedures, see the *PISA 2015 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 science, reading or mathematics instructor of any students in the sessions he or she would conduct 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 reviewing and updating the Student Tracking Form; completing the Session Attendance Form, which is designed to record students' attendance and instruments allocation; completing the Session Report Form, which is designed to summarise session times, any disturbance to the session, etc.; ensuring that the number of test booklets and questionnaires collected from students tallied with the number sent to the school (paper-based assessment countries) or ensuring that the number of USB sticks used for the assessment were accounted for (computer-based assessment countries); and sending the school questionnaire, student questionnaires, parent and teacher questionnaires (if applicable), and all test materials (both completed and not completed) to the national centre after the testing.

The PISA Consortium responsible for overseeing survey operations implemented all phases of the PISA Quality Monitor (PQM) process: interviewing and hiring PQM candidates in each of the countries, organising their training, selecting the schools to visit, and collecting information from the PQM visits. PQMs are independent contractors located in participating countries who are hired by the international survey operations contractor. They visit a sample of schools to observe test administration and to record the implementation of the documented field-operations procedures in the main survey.

Typically, two or three PQMs were hired for each country, and they visited an average of 15 schools in each country. If there were adjudicated regions in a country, it was usually necessary to hire additional PQMs, as a minimum of five schools were observed in adjudicated regions.

All quality-assurance data collected throughout the PISA 2015 assessment were entered and collated in a central data-adjudication database on the quality of field operations, printing, translation, school and student sampling, and coding.

Comprehensive reports were then generated for the PISA Adjudication Group. This group was formed by the Technical Advisory Group and the Sampling Referee. Its role is to review the adjudication database and reports to recommend adequate treatment to preserve the quality of PISA data. For further information, see the *PISA 2015 Technical Report* (OECD, forthcoming).

The results of adjudication and subsequent further examinations showed that the PISA Technical Standards were met in all countries and economies that participated in PISA 2015 except for those countries listed below:

- In Albania, the PISA assessment was conducted in accordance with the operational standards and guidelines of the OECD. However, because of the ways in which the data were captured, it was not possible to match the data in the test with the data from the student questionnaire. As a result, Albania cannot be included in analyses that relate students' responses from the questionnaires to the test results.
- In Argentina, the PISA assessment was conducted in accordance with the operational standards and guidelines of the OECD. However, there was a significant decline in the proportion of 15-year-olds who were covered by the test, both in absolute and relative numbers. There had been a re-structuring of Argentina's secondary schools, except for those in the adjudicated region of Ciudad Autónoma de Buenos Aires, which is likely to have affected the coverage of eligible schools listed in the sampling frame. As a result, Argentina's results may not be comparable to those of other countries or to results for Argentina from previous years.
- In Kazakhstan, the national coders were found to be lenient in marking. Consequently, the human-coded items did not meet PISA standards and were excluded from the international data. Since human-coded items form an important part of the constructs that are tested by PISA, the exclusion of these items resulted in a significantly smaller coverage of the PISA test. As a result, Kazakhstan's results may not be comparable to those of other countries or to results for Kazakhstan from previous years.
- In Malaysia, the PISA assessment was conducted in accordance with the operational standards and guidelines of the OECD. However, the weighted response rate among the initially sampled Malaysian schools (51%) falls well short of the standard PISA response rate of 85%. Therefore, the results may not be comparable to those of other countries or to results for Malaysia from previous years.

Reference

OECD (forthcoming), *PISA 2015 Technical Report*, OECD Publishing, Paris.

ANNEX A5

CHANGES IN THE ADMINISTRATION AND SCALING OF PISA 2015 AND IMPLICATIONS FOR TRENDS ANALYSES

Available on line only.

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

ANNEX A6
GUIDELINES AND CAVEATS ABOUT INTERPRETING THE RESULTS

Interpreting the data from students, parents and schools

PISA 2015 asked students and school principals to answer questions about the learning environment and organisation of schools, and the social and economic contexts in which learning takes place. Information based on their responses has been weighted so that it reflects the number of 15-year-old students enrolled in grade 7 or above. These are self-reports rather than external observations and may be influenced by cultural differences in how individuals respond. For example, individual students in the same classroom may perceive and report classroom situations in different ways, or respondents may provide responses that are considered to be more socially desirable or acceptable than others.

In addition to the general limitation of self-reported data, there are other limitations, particularly those concerning the information collected from principals, that should be taken into account when interpreting the data:

- On average across OECD countries, 268 principals were surveyed, but in 10 countries and economies, fewer than 150 principals were surveyed, and in Ciudad Autónoma de Buenos Aires (Argentina), Luxembourg, Macao (China), Malta and Montenegro, fewer than 100 principals were surveyed (see Table A7.1 from Annex A7 of Volume II). Although principals can provide information about their schools, generalising from a single source of information for each school is not straightforward. Also, principals' perceptions may not be the most appropriate sources of some information related to teachers, such as teachers' morale and commitment.
- Students' attitudes towards learning and their performance in each subject depend on many factors, including all the education that they have acquired in previous years and their experiences outside the school setting. In most cases, 15-year-old students have been in their current school for only two or three years. The learning environment examined by PISA may therefore only partially reflect the learning environment that shaped students' experiences in education earlier in their school careers. To the extent that students' current learning environment differs from that of their earlier school years, the contextual data collected by PISA are an imperfect proxy for students' cumulative learning environments.
- In some countries and economies, the definition of the school in which students are taught is not straightforward because schools vary in the level and purpose of education. For example, in some countries and economies, subunits within schools (e.g. study programmes, shifts and campuses) were sampled instead of schools as administrative units. See Annex A2 for further information.
- The age-based sampling followed in PISA means that, in some education systems, students are not always representative of their schools. Interpreting differences between schools correctly therefore requires specific knowledge about how school systems are structured.

Despite these caveats, information from the school questionnaire provides unique insights into the ways in which national and subnational authorities seek to realise their education objectives.

Schooling and school effects

In using results from non-experimental data on school performance, such as the PISA Database, it is important to bear in mind the distinction between school effects and the effects of schooling, particularly when interpreting the modest association between factors such as school resources, policies and institutional characteristics and student performance. School effects are education researchers' shorthand for the effect on academic performance of attending one school or another, usually schools that differ in resources or policies and institutional characteristics. Where schools and school systems do not vary in fundamental ways, the school effect can be modest. Nevertheless, modest school effects should not be confused with a lack of an effect of schooling (the influence on performance of not being schooled compared with being schooled).

Interpreting correlations

A correlation is a simple statistic that measures the degree to which two variables are associated with each other, but does not prove causality between the two.

Interpreting results before and after accounting for socio-economic status

When examining the relationship between education outcomes and resources, policies and practices within school systems, this volume takes into account the socio-economic differences among students and schools. The advantage of doing this lies in comparing similar entities, namely students and schools with similar socio-economic profiles. At the same time, there is a risk that such adjusted comparisons underestimate the strength of the relationship between student performance and resources, policies and practices, since most of the differences in performance are often attributable to both policies and socio-economic status.

Conversely, analyses that do not take socio-economic status into account can overstate the relationship between student performance and resources, policies and practices, as the level of resources and the kinds of policies adopted may also relate to the socio-economic profile of students, schools and countries and economies. At the same time, analyses without adjustments may paint a more realistic picture of the schools that parents choose for their children. They may also provide more information for other stakeholders who are interested in the overall performance of students, schools and systems, including any effects that may be related to the socio-economic profile of schools and systems. For example, parents may be primarily interested in a school's absolute performance standards, even if a school's higher achievement record stems partially from the fact that the school has a larger proportion of advantaged students.

Interpreting the results by school characteristics

When presenting the results by the socio-economic profile of schools, the location of schools, the type of school or the education level, the number of students and schools in each subsample has to meet the PISA reporting requirements of at least 30 students and 5 schools. Even when these reporting requirements are met, the reader should interpret the results cautiously when the number of students or schools is just above the threshold. Table A7.1 (OECD, 2016) shows the unweighted number of students and schools by school characteristics in the PISA sample so that the reader can interpret the results appropriately.

Interpreting odds ratios

An odds ratio indicates the degree to which an explanatory variable is associated with a categorical outcome variable with two categories (e.g. yes/no) or more than two categories. An odds ratio below one denotes a negative association; an odds ratio above one indicates a positive association; and an odds ratio of one means that there is no association.

Imagine that the association between being a boy and having repeated a grade is being analysed, the following odds ratios would be interpreted as:

- 0.2 > Boys are five times less likely to have repeated a grade than girls.
- 0.5 > Boys are half as likely to have repeated a grade as girls.
- 0.9 > Boys are 10% less likely to have repeated a grade than girls.
- 1.0 > Boys and girls are equally likely to have repeated a grade.
- 1.1 > Boys are 10% more likely to have repeated a grade than girls.
- 2.0 > Boys are twice more likely to have repeated a grade than girls.
- 5.0 > Boys are five times more likely to have repeated a grade than girls.

Reference

OECD (2016), *PISA 2015 Results (Volume II): Policies and Practices for Successful Schools*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264267510-en.