



# Annex A

## TECHNICAL NOTES ON ANALYSES IN THIS REPORT



## Sources of data

PISA datasets are the main source of data used in this report. However, in some chapters, additional sources of data are used. Data from the Survey of Adult Skills (PIAAC) (OECD, 2016<sup>[1]</sup>) are used in Chapters 2 and 5; data from the Barro-Lee Educational Attainment Dataset (Barro and Lee, 2013<sup>[2]</sup>) are used in Chapter 2; and data from the Trends in International Mathematics and Science Study (TIMSS) (IEA, 1997<sup>[3]</sup>) are also used in Chapter 2. Additionally, Chapter 5 uses longitudinal data accessed by the OECD Secretariat, in its raw form or through collaboration with national researchers, from five countries: Australia, Canada, Denmark, Switzerland and the United States. For a detailed description of the longitudinal data used in Chapter 5, see Box 5.1.

## Explanation of the PISA indices used in the report

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

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, 2017<sup>[4]</sup>) provides an in-depth description of this conceptual framework. Structural equation modelling was used to confirm the theoretically expected behaviour of most 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 PISA indices and details on the methods, see the *PISA 2015 Technical Report* (OECD, 2017<sup>[5]</sup>).

There are two types of indices used in this volume: simple indices and scale indices. Simple indices are variables that are constructed through the arithmetic transformation or recoding of one or more items in exactly the same way across assessments. Scale indices are 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, 1985<sup>[6]</sup>). For details on how each scale index was constructed, see the *PISA 2015 Technical Report* (OECD, 2017<sup>[5]</sup>).

In addition to the 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 a prefix of “ST” for items in the student questionnaire, “SC” for items in the school questionnaire, “PA” for items from the parent questionnaire, and “IC” for items from the ICT questionnaire. All the context questionnaires as well as the PISA international database, including all variables, are available through [www.oecd.org/pisa](http://www.oecd.org/pisa).

## Student-level simple indices

### *Students’ age*

The age of a student (AGE) was calculated as the difference between the year and month of testing and the year and month of the student’s birth. Data on students’ age were obtained from both the questionnaire (ST003) and 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 to questions ST005, ST006, ST007 and ST008 regarding their parents’ education were classified using ISCED 1997 (OECD, 1999<sup>[11]</sup>). Indices on parents’ 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 <ISCED level 4> (non-tertiary post-secondary), (5) <ISCED level 5B> (vocational tertiary) and (6) <ISCED level 5A> and <ISCED level 6> (theoretically-oriented tertiary and post-graduate). Indices with these categories were constructed for each student's mother (MISCED) and father (FISCED). In addition, the index of parents' highest level of education (HISCED) corresponds to the higher ISCED level of either parent. The index of parents' highest level of education was also recorded by the 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, 2017<sup>[5]</sup>).

### **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 (International Standard Classification of Occupations) codes (International Labour Office, 2012<sup>[7]</sup>) and then mapped to the International Socio-Economic Index of occupational status (ISEI) (Ganzeboom and Treiman, 2003<sup>[8]</sup>). In PISA 2015, as in PISA 2012, the new ISCO codes and ISEI index (2008 version) were used rather than the 1998 versions that had been applied in the previous four cycles of PISA (Ganzeboom and Treiman, 2010<sup>[9]</sup>). 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.

### **Students' expected occupational status**

In the PISA 2015 student questionnaire, as in the PISA 2006 questionnaire, students were asked: "What kind of job do you expect to have when you are about 30 years old?" This was an open question, meaning that no response categories were provided and students were able to answer freely using their own words. Responses were coded to four-digit ISCO codes and then mapped to the ISEI (International Socio-Economic Index of occupational status) index (Ganzeboom and Treiman, 2010<sup>[9]</sup>). Higher scores in the ISEI index indicate higher occupational status.

### **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, including 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 at the age of 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 includes 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 career 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. *Science and engineering professionals*: All science and engineering professionals (submajor group 21), except product and garment designers (2163) and 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 communication technology professionals (submajor group 25).
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).

For further details about this index, including how this classification compares to existing classifications, see Annex A1 in *PISA 2015 Results (Volume I)* (OECD, 2016<sub>[10]</sub>).

### **International academic resilience**

The index of international academic resilience is a simple index that takes the value of one if a student meets two conditions and takes the value of zero otherwise. A student is classified as “internationally resilient” if the student is in the bottom quarter of the PISA index of economic, social and cultural status (ESCS) in the country/economy of assessment; and if the student scored in the top quarter of performance in science among all students participating in PISA, after accounting for socio-economic background.

### **National academic resilience**

The index of national academic resilience is a simple index that takes the value of one if a student meets two conditions and takes the value of zero otherwise. A student is classified as “nationally resilient” if the student is in the bottom quarter of the ESCS index in the country/economy of assessment, and if the student scored in the top quarter of performance in science among students in his or her own country.

### **Core-skills academic resilience**

The index of core-skills academic resilience is a simple index that takes the value of one if a student meets two conditions and takes the value of zero otherwise. A student is classified as “core-skills resilient” if the student is in the bottom quarter of the ESCS index in the country/economy of assessment, and if the student scored at or above PISA proficiency Level 3 in science, reading and mathematics.

### **Social and emotional resilience**

The index of social and emotional resilience is a simple index that takes the value of one if a student meets four conditions and takes the value of zero otherwise. A student is classified as “socially and



emotionally resilient” if she or he meets the following four criteria: i) the student is in the bottom quarter of the ESCS index in the country/economy of assessment; ii) the student rated her or his life satisfaction with a value from seven to ten, on a scale from zero to ten, with zero meaning “not at all satisfied” and ten meaning “completely satisfied” (questionnaire item ST016 was: “Overall, how satisfied are you with your life as a whole these days?”); iii) the student disagreed with the following statement: “I feel like an outsider (or left out of things) at school”; and iv) the student disagreed with the following statement: “Even when I am well prepared for a test I feel very anxious”.

### **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<sup>[11]</sup>). 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.

### **Immigrant background**

The PISA database contains three country-specific variables relating to the country of birth of the student, their mother and their 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 the country of assessment and (2) other. The index of immigrant background (IMMIG) was calculated from these variables with the following categories: (0) non-immigrant students (those students who had at least one parent born in the country), and (1) first- and second-generation immigrant students (those born outside the country of assessment and whose parent[s] were also born in another country; and those born in the country of assessment but whose parent[s] were born in another country). Students with missing responses for either the student or for both parents were assigned missing values for this variable.

## **Student-level scale indices**

### **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 occupational 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.

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 (ESCS) was derived from a principal component analysis of standardised variables (each variable has an OECD mean of zero and a standard deviation of one), taking the factor scores for the first principal component as measures of this index. 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 was 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 and economies.

For more detailed information on how the ESCS was constructed, please refer to the *PISA 2015 Technical Report* (OECD, 2017<sup>[5]</sup>).

### **Achievement motivation**

The index of achievement motivation (MOTIVAT) was constructed from students' responses to a new question developed for PISA 2015 (ST119). Students were asked to report their agreement ("strongly disagree"; "disagree"; "agree"; or "strongly agree") 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"; and "I want to be one of the best students in my class". Higher values indicate that students have greater achievement motivation.

### **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 to rate how well they would perform ("I could do this easily"; "I could do this with a bit of effort"; "I would struggle to do this on my own"; or "I couldn't do this") 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 in 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.

### **Disciplinary climate**

The index of disciplinary climate in science classes (DISCLSCI) was constructed from students' reports on how often ("every lesson"; "most lessons"; "some lessons"; or "never or hardly ever") the following happened in their science lessons (ST097): "students don't listen to what the teacher says"; "there is



noise and disorder”; “the teacher has to wait a long time for students to quiet down”; “students cannot work well”; and “students don’t start working for a long time after the lesson begins”.

### ***Index of effort and perseverance (EFFPER)***

The PISA 2000 index of effort and perseverance was derived from students’ reports on how often (“almost never”; “sometimes”; “often”; or “almost always”) the following statements applied to the student: “when studying, I work as hard as possible”; “when studying, I keep working even if the material is difficult”; “when studying, I try to do my best to acquire the knowledge and skills taught”; and “when studying, I put forth my best effort”. Scale scores are standardised Warm estimates where positive values indicate greater frequency and negative values indicate lower frequency of using effort and perseverance as a learning strategy.

### ***Index of cultural communication (CULTCOM)***

The PISA 2000 index of cultural communication was derived from students’ reports on the frequency (“never or hardly ever”; “a few times a year”; “about once a month”; “several times a month”; or “several times a week”) with which their parents (or guardians) engaged with them in: discussing books, films or television programmes; and listening to classical music. Scale scores are standardised Warm estimates where positive values indicate a greater frequency and negative values indicate a lower frequency of cultural communication.

### ***Index of cultural activities (CULTACTV)***

The PISA 2000 index of cultural activities was derived from students’ reports on how often during the preceding year (“never or hardly ever”; “once or twice a year”; “about three or four times a year”; or “more than four times a year”) they had: visited a museum or art gallery; attended an opera, ballet or classical symphony concert; or watched live theatre. Scale scores are standardised Warm estimates where positive values indicate a greater frequency and negative values indicate a lower frequency of participating in cultural activities during the year.

### ***Teacher-directed science instruction***

The index of teacher-directed science instruction (TDTEACH) was constructed from students’ reports on how often (“never or almost never”; “some lessons”; “many lessons”; or “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”; and “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”; or “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”; and “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”; or “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”; and “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”; or “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 <school 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 <broad science> concepts to our lives”; and “students are asked to do an investigation to test ideas”.

### ***Index of interest in reading (INTREA)***

The PISA 2000 index of interest in reading was derived from students’ level of agreement (“disagree”; “disagree somewhat”; “agree somewhat”; or “agree”) with the following three statements: “because reading is fun, I wouldn’t want to give it up”; “I read in my spare time”; and “when I read, I sometimes get totally absorbed”. Scale scores are standardised Warm estimates where positive values indicate higher levels and negative values indicate lower levels of interest in reading.

### ***Index of engagement in reading (JOYREAD)***

The PISA 2000 index of engagement in reading was derived from students’ level of agreement (“strongly disagree”; “disagree”; “agree”; or “strongly agree”) with the following statements: “I read only if I have to”; “reading is one of my favourite hobbies”; “I like talking about books with other people”; “I find it hard to finish books”; “I feel happy if I receive a book as a present”; “for me, reading is a waste of time”; “I enjoy going to a bookstore or library”; “I read only to get information that I need”; and “I cannot sit still and read for more than a few minutes”. Scale scores are standardised Warm estimates where positive values indicate more positive attitudes and negative values indicate less positive attitudes towards reading.

## **School-level simple indices**

### ***Science-specific resources***

The index of science-specific resources (SCIRES) was constructed using principals’ responses to a series of statements about the school science department. It was constructed by summing up principals’ answers to the following eight statements in SC059: “compared to other departments, our schools’ <school science department> is well equipped”; “if we ever have some extra funding, a big share goes into improvement of our <school science> teaching”; “<school science> teachers are among our best educated staff members”; “compared to similar schools, we have a well-equipped laboratory”; “the material for hands-on activities in <school science> is in good shape”; “we have enough laboratory material that all courses can regularly use it”; “we have extra laboratory staff that helps support <school science> teaching”; and “our school spends extra money on up-to-date <school science> equipment”.

### ***Class size***

The average class size (CLSIZE) is derived from one of nine possible categories in question SC003, ranging from “15 students or fewer” to “more than 50 students”.





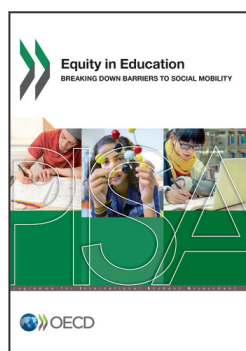
## School-level scale indices

### **Student behaviour hindering learning**

The index of student behaviour hindering learning (STUBEHA) was constructed from school principals' reports on the extent ("not at all"; "very little"; "to some extent"; or "a lot") to which they think that student learning in their schools is hindered by the following phenomena (SC061): student truancy; students skipping classes; students lacking respect for teachers; students using alcohol or illegal drugs; and students intimidating or bullying other students. The responses were combined to create an index such that, across OECD countries, the mean is zero and the standard deviation is one.

## References

- Barro, R. and J. Lee (2013), "A new data set of educational attainment in the world, 1950–2010", *Journal of Development Economics*, Vol. 104, pp. 184–198, <http://dx.doi.org/10.1016/J.JDEVECO.2012.10.001>. [2]
- Ganzeboom, H. and D. Treiman (2010), *Occupational Status Measures for the New International Standard Classification of Occupations ISCO-08; With a Discussion of the New Classifications*, <http://www.harryganzeboom.nl/isol/isol2010c2-ganzeboom.pdf> (accessed on 02 August 2018). [9]
- Ganzeboom, H. and D. Treiman (2003), "Three Internationally Standardised Measures for Comparative Research on Occupational Status", in *Advances in Cross-National Comparison*, Springer US, Boston, MA, [http://dx.doi.org/10.1007/978-1-4419-9186-7\\_9](http://dx.doi.org/10.1007/978-1-4419-9186-7_9). [8]
- IEA (1997), *Mathematics in the Primary School Years: IEA's Third International Mathematics and Science Report*, International Association for the Evaluation of Educational Achievement, Boston, <https://timssandpirls.bc.edu/timss1995i/MathA.html> (accessed on 02 August 2018). [3]
- International Labour Office (2012), *International Standard Classification of Occupations, ISCO-08*, <http://www.ilo.org/public/english/bureau/stat/isco/docs/publication08.pdf> (accessed on 02 August 2018). [7]
- OECD (2017), *PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematic, Financial Literacy and Collaborative Problem Solving*, PISA, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264281820-en>. [4]
- OECD (2017), *PISA 2015 Technical Report*, OECD, Paris, <http://www.oecd.org/pisa/sitedocument/PISA-2015-technical-report-final.pdf> (accessed on 02 August 2018). [5]
- OECD (2016), *PISA 2015 Results (Volume I): Excellence and Equity in Education*, PISA, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264266490-en>. [10]
- OECD (2016), *The Survey of Adult Skills: Reader's Companion, Second Edition*, OECD, Paris, <http://dx.doi.org/10.1787/9789264258075-en> (accessed on 02 August 2018). [1]
- OECD (1999), *Classifying Educational Programmes: Manual for ISCED-97 Implementation in OECD Countries*, <http://www.oecd.org/education/skills-beyond-school/1962350.pdf> (accessed on 02 August 2018). [11]
- Warm, T. (1985), "Weighted maximum likelihood estimation of ability in item response theory with tests of finite length", *Psychometrika*, Vol. 54/3, pp.427–450, <https://dx.doi.org/10.1007/BF02294627>. [6]



From:

## Equity in Education

Breaking Down Barriers to Social Mobility

Access the complete publication at:

<https://doi.org/10.1787/9789264073234-en>

### Please cite this chapter as:

OECD (2018), “Technical notes on analyses in this report”, in *Equity in Education: Breaking Down Barriers to Social Mobility*, OECD Publishing, Paris.

DOI: <https://doi.org/10.1787/9789264073234-9-en>

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

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgment of OECD as source and copyright owner is given. All requests for public or commercial use and translation rights should be submitted to [rights@oecd.org](mailto:rights@oecd.org). Requests for permission to photocopy portions of this material for public or commercial use shall be addressed directly to the Copyright Clearance Center (CCC) at [info@copyright.com](mailto:info@copyright.com) or the Centre français d'exploitation du droit de copie (CFC) at [contact@cfcopies.com](mailto:contact@cfcopies.com).