

ANNEX A1

Construction of indices

EXPLANATION OF INDICES

This section explains the indices derived from the PISA 2018 student, school, parent and ICT 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 on the basis of theoretical considerations and previous research. The *PISA 2018 Assessment and Analytical Framework* (OECD, 2019_[1]) provides an in-depth description of this conceptual framework. Item response theory (IRT) modelling was used to confirm the theoretically expected behaviour of the indices and to validate their comparability across countries. For a detailed description of the methods, see the section “Cross-country comparability of scaled indices” in this chapter, and the *PISA 2018 Technical Report* (OECD, forthcoming_[2]).

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.

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_[1]). For details on how each scale index was constructed, see the *PISA 2018 Technical Report* (OECD, forthcoming_[2]). In general, the scaling was done in two stages:

- The item parameters were estimated based on all students from equally weighted 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 some trend indices, a common calibration linking procedure was used: countries/economies that participated in both PISA 2009 and PISA 2018 contributed both samples to the calibration of item parameters; each cycle and, within each cycle, each country/economy contributed equally to the estimation.
- 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 were given equal weight in the standardisation process).

Sequential codes were assigned to the different response categories of the questions in the sequence in which the latter appeared in the student, school or parent questionnaire. Where indicated in this section, these codes were inverted for the purpose of constructing indices or scales. Negative values in an index do not necessarily imply that students responded negatively to the underlying questions. A negative value merely indicates that a respondent answered less positively than other respondents did on average across OECD countries. Likewise, a positive value in an index indicates that a respondent answered more favourably, or more positively, on average, than other 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. All the context questionnaires, and the PISA international database, including all variables, are available at www.oecd.org/pisa.

STUDENT-LEVEL SIMPLE INDICES

Immigrant background

Information was collected on the country of birth of the students and their parents. Included in the database are three country-specific variables related to the country of birth of the student, mother and father (ST019). The variables are binary and indicate whether the student, mother and father were born in the country of assessment or elsewhere. The index on immigrant background (IMMIG) is calculated from these variables, and has the following categories: (1) native students (those students

with at least one parent who was born in the country); (2) second-generation students (those students born in the country of assessment but whose parents were born in another country); and (3) first-generation students (those students born outside the country of assessment and whose parents were also born outside the country of assessment). Students with missing responses for either the student or for both parents were given 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, provided that the student had not repeated a grade in any of the other ISCED levels. The index was assigned a missing value if none of the three categories were ticked for any of the three ISCED levels.

Study programme indices

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 (ST002). 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 1997). 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); and programme orientation (ISCEDO) indicates whether the programme’s curricular content was general, pre-vocational or vocational.

Early childhood education and care

Questions ST125 and ST126 measure the starting age in ISCED 1 and ISCED 0. The indicator DURECEC is built as the difference of ST126 and ST125 plus the value of “2” to indicate the number of years a student spent in early childhood education and care.

Learning time

Learning time in the test language (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).

Expected occupational status

As in previous cycles of PISA, students were asked to report their expected occupation at age 30 and a description of this job (ST114). The responses were coded to four-digit ISCO codes (OCOD3) and then mapped to the ISEI index (Ganzeboom and Treiman, 2003_[11]). Recoding of ISCO codes into ISEI index results in scores for the student’s expected occupational status (BSMJ), where higher scores of ISEI indicate higher levels of expected occupational status.

STUDENT-LEVEL SCALE INDICES

Sense of belonging

The index of sense of belonging (BELONG) was constructed using students’ responses to a trend question about their sense of belonging at school. Students were asked whether they agree (“strongly disagree”, “disagree”, “agree”, “strongly agree”) with the following school-related 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”; and “I feel lonely at school”. Positive values in this scale mean that students reported a greater sense of belonging at school than did the average student across OECD countries.

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 occupational status (HISEI), and home possessions (HOMEPOS), including books in the home.

Parents’ highest level of education

Students’ responses to questions ST005, ST006, ST007 and ST008 regarding their parents’ education were classified using ISCED 1997 (OECD, 1999_[5]). 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), and the index of highest education level of parents (HISCED) corresponded 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). In PISA 2018, to avoid issues related to the misreporting of parental education by students, students' answers about post-secondary qualifications were considered only for those students who reported their parents' highest level of schooling to be at least lower secondary education. The conversion from ISCED levels to year of education is common to all countries. This international conversion was determined by using the modal years of education across countries for each ISCED level. The correspondence is available in the *PISA 2018 Technical Report* (OECD, forthcoming^[2]).

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^[1]). In PISA 2018, as in PISA 2015, 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^[2]). 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. In PISA 2018, in order to reduce missing values, an ISEI value of 17 (equivalent to the ISEI value for ISCO code 9000, corresponding to the major group "Elementary Occupations") was attributed to pseudo-ISCO codes 9701, 9702 and 9703 ("Doing housework, bringing up children", "Learning, studying", "Retired, pensioner, on unemployment benefits").

Household possessions

In PISA 2018, 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, ST013). HOMEPOS is a summary index of all household and possession items (ST011, ST012 and ST013).

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.

In previous cycles, 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 0 and a standard deviation of 1), taking the factor scores for the first principal component as measures of the PISA index of economic, social and cultural status. In PISA 2018, ESCS was computed by attributing equal weight to the three standardised components. As in PISA 2015, the three components were standardised across all countries and economies (both OECD and partner countries/economies), with each country/economy contributing equally (in cycles prior to 2015, the standardisation and principal component analysis was based on OECD countries only). As in every previous cycle, the final ESCS variable was transformed, with 0 the score of an average OECD student and 1 the standard deviation across equally weighted OECD countries.

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 to make decisions concerning its affairs (Question SC013). Public schools are managed directly or indirectly by a public education authority, government agency or governing board appointed by government or elected by public franchise. Private schools are managed directly or indirectly by a non-government organisation, such as a church, trade union, business or other private institution. In some countries and economies, such as Ireland, the information from SC013 was combined with administrative data to determine whether the school is privately or publicly managed.

Socio-economic profile of the schools

Advantaged and disadvantaged schools are defined in terms of the socio-economic profile of schools. All schools in each PISA-participating education system are ranked according to their average PISA index of economic, social and cultural status

(ESCS) and then divided into four groups with approximately an equal number of students (quarters). Schools in the bottom quarter are referred to as “socio-economically disadvantaged schools”; and schools in the top quarter are referred to as “socio-economically advantaged schools”.

School size

The index of school size (SCHSIZE) contains the total enrolment at school. It is based on the enrolment data provided by the school principal, summing up the number of girls and boys at a school (SC002). This index was calculated in 2018 and in all previous cycles.

Class size

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

Availability of computers

School principals were asked to report the number of computers available at school (SC004). The index of availability of computers (RATCMP1) is the ratio of computers available to 15-year-olds for educational purposes to the total number of students in the modal grade for 15-year-olds. The index RATCMP2 was calculated as the ratio of number of computers available to 15-year-olds for educational purposes to the number of these computers that were connected to the Internet.

Quantity of teaching staff at school

Principals were asked to report the total number of teachers at their school (TOTAT).

Extracurricular activities at school

School principals were asked to report what extracurricular activities their schools offered to 15-year-old students (SC053). The index of creative extracurricular activities at school (CREACTIV) was computed as the total number of the following activities that occurred at school: i) band, orchestra or choir; ii) school play or school musical; and iii) art club or art activities.

SCHOOL-LEVEL SCALE INDICES

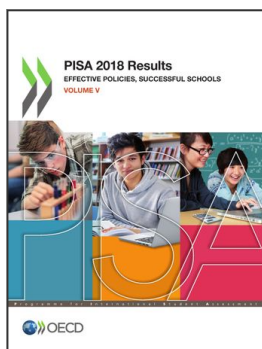
Indices included in earlier assessments

School resources

As in PISA 2015 and 2012, PISA 2018 included an eight-item question (SC017) about school resources, measuring school principals' perceptions of potential factors hindering instruction at school (“Is your school's capacity to provide instruction hindered by any of the following issues?”). The four response categories were: “not at all”, “very little”, “to some extent”, “a lot”. A similar question was used in previous cycles, but items were reduced and reworded for 2012 focusing on two derived variables. The index of staff shortage (STAFFSHORT) was derived from the first four items: a lack of teaching staff; inadequate or poorly qualified teaching staff; a lack of assisting staff; inadequate or poorly qualified assisting staff. The index of educational material shortage (EDUSHORT) was derived from the second set of four items: a lack of educational material; inadequate or poor quality educational material; a lack of physical infrastructure; inadequate or poor quality physical infrastructure. Positive values in this index mean that principals viewed the amount and/or quality of the human or educational resources in their schools as an obstacle to providing instruction to a greater extent than the OECD average.

References

- Ganzeboom, H.** (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”. [6]
- Ganzeboom, H.** and **D. Treiman** (2003), “Three internationally standardised measures for comparative research on occupational status”, *Advances in cross-cultural comparison*. [4]
- OECD** (2019), *PISA 2018 Assessment and Analytical Framework*, OECD Publishing, Paris, <https://dx.doi.org/10.1787/b25efab8-en>. [1]
- OECD** (1999), *Classifying Educational Programmes: Manual for ISCED-97 Implementation in OECD Countries*, OECD Publishing, Paris, <http://www.oecd.org/education/1841854.pdf> (accessed on 28 October 2019). [5]
- OECD** (forthcoming), *PISA 2018 Technical Report*, OECD Publishing, Paris. [2]
- Warm, T.** (1989), “Weighted likelihood estimation of ability in item response theory”, *Psychometrika*, Vol. 54/3, pp. 427-450, <https://dx.doi.org/10.1007/BF02294627>. [3]



From:
PISA 2018 Results (Volume V)
Effective Policies, Successful Schools

Access the complete publication at:

<https://doi.org/10.1787/ca768d40-en>

Please cite this chapter as:

OECD (2020), "Construction of indices", in *PISA 2018 Results (Volume V): Effective Policies, Successful Schools*, OECD Publishing, Paris.

DOI: <https://doi.org/10.1787/84a683c1-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, as well as any data and 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. Extracts from publications may be subject to additional disclaimers, which are set out in the complete version of the publication, available at the link provided.

The use of this work, whether digital or print, is governed by the Terms and Conditions to be found at <http://www.oecd.org/termsandconditions>.